dup RS

Vol. 45 FEBRUARY, 1954 No. 2

The Journal of Educational Psychology

CONTENTS

Educational Therapy 65
BRUCE McCARTY

Who Is Penalized by the Penality for Guessing? 81
ALEX C. SHERRIFFS AND DONALD S. BOOMER

Topographical Orientation in Naval Aviation Cadets 91
BRANT CLARK AND R. DANIEL MALONE

Student Differences in Competency in Educational Psychology Prior to Formal Training 110
WILLIAM W. LYNCH, JR.

Book Reviews 121

Published monthly except June to September
$7.00 a year in the United States; Canada, $7.20; other countries, $7.40.
Single issues $1.20

WARWICK & YORK, INC.
BALTIMORE 2, MD.

Entered as Second Class Matter Nov. 15, 1921, at the Post Office at Baltimore, Md. under the Act of March 3, 1879.

THE JOURNAL OF Educational Psychology

Established 1910

BOARD OF EDITORS*

STEPHEN M. COREY, *Chairman*
Teachers College
Columbia University

J. B. STROUD
State University of Iowa

WM. CLARK TROW
University of Michigan

*Appointed by the American Psychological Association.

EDITORIAL CONSULTANTS

JACK W. DUNLAP
KARL J. HOLZINER
HAROLD E. JONES
H. H. REMMERS
PERCIVAL M. SYMONDS
MILES A. TINKER
ALEXANDER G. WESMAN
PAUL A. WITTY

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY is devoted primarily to the scientific study of problems of learning, teaching, and measurement of the psychological development of the individual. The JOURNAL will contain articles on the following subjects: the psychology of school subjects; experimental studies of learning; the development of interests, attitudes, and personality, particularly as related to school adjustment; emotion, motivation, and character; mental development and methods. This last will include tests, statistical techniques, and research techniques in cross-sectional and developmental studies.

INFORMATION FOR CONTRIBUTORS

Manuscripts and communications regarding editorial matters should be addressed to the Chairman of the Board of Editors, Stephen M. Corey, Teachers College, Columbia University, New York 27, N. Y.

THE JOURNAL has set regulations regarding content and style of material published, and these should be observed in the preparation of manuscripts to be submitted.

Tables and graphs.—In keeping with the practice of most professional periodicals of like character, THE JOURNAL now requires authors to bear part of the added cost resulting from the use of tables, formulas, and graphs. Prospective contributors may obtain a detailed statement as to

(Continued on Inside Back Cover)

THE JOURNAL OF

EDUCATIONAL PSYCHOLOGY

Volume 45 February, 1954 Number 2

EDUCATIONAL THERAPY

BRUCE McCARTY

Waco, Texas

This is a report on the Experimental Educational Therapy program which was conducted at the Austin State Hospital from March to the end of August of 1951. It was cosponsored by the Hospital and the Hogg Foundation for Mental Health.

RATIONALE

Mental institutions have been thought of historically as places where unfortunates were kept away from civilization for their mutual protection. Not until comparatively recently has it been thought possible to restore mental patients to a healthy active rôle in society, and with the dawn of this concept, the name of the institution was changed from 'asylum' to 'hospital'. Sad to relate, the practice has not kept pace with the theory. In many instances the professed intention of relieving mental illness and restoring the individual has taken a back seat to administrative convenience which is often suffused with the file-and-forget philosophy. More often than not, various therapies which are in use are merely peripheral measures designed more to occupy time than to ameliorate aberrances.

It seems likely that even though the etiology of mental disorders is primarily emotional, the patient's preoccupation with these personality problems interferes with his intellectual accomplishments and this in turn causes him to fall behind in his own rate of development—in what Havighurst calls 'developmental tasks'. Though these are frequently thought of as being problems of a strictly academic nature, such is only part of the story; sundry items such as the acquisitions of graceful interpersonal relationships, advancement in current activities (whether educational, occupational or

personal in nature) and just generally keeping up with the-here-and-the-now are no less developmental tasks which must be mastered in their proper sequence to form stepping stones toward future life manipulation. The befoulment of this orderly sequence of developmental progress in turn causes an emotional problem of its own. As the individual sees himself being left behind by the onrushing world his urge for inferiority-avoidance reaches near panic proportions and the vicious cycle thus created is not materially relieved by therapies designed to waste yet more time.

Clearly, if these two facets of the problem are simultaneously operative they must be attacked simultaneously in order to render therapy effective. There must be (1) the traditional psycho-environmental attack on the emotional problem: environmental in order to release the patient from the tension causing rigidity of his old milieu, and psychotherapeutic to rebuild an ego substantial enough to withstand the stresses and strains of life.

Concurrently, there must be provided (2) the opportunity to bring expectations and accomplishments into line. It is hard to see how there could be any successful and permanent rebuilding of the ego with only the materials which were present from the incipient stages of the morbid condition. These were evidently inadequate as a bulwark once, and there is little reason to suppose them adequate as time and tide go on apace. The addition of tools for living is essential, whether academic, professional and/or avocational; and in its widest sense we call this increment Educational Therapy.

CLASS PLAN

The idea of Educational Therapy for the younger patients of the Texas State Hospital was primarily a pilot study for assessing the effects of education as a means of therapy for the students, and also the effect of these abnormal students on education.

The curriculum was academic- rather than vocational-centered. Of course this does not mean that it was built exclusively around the three 'R's'. A typing program, for instance, radio course and art classes were pointed in the direction of specific trades, but the emphasis was on the traditional subjects.

The school was set up in the former office of the beauty shop. This room was particularly well suited for the purpose in that it was light, modern-looking, well ventilated and out of the heavy

traffic flow. The walls were a combination of stained wood and plaster, the latter part of which was used as a movie screen. One of the two long tables contained our book collection and three archaic typewriters (one of which was in such condition that it served only in the capacity of a bookend). The other table was work space.

Books on arithmetic, art, astronomy, chemistry, civics, English, geography, geometry, health, history, physics, radio, spelling and trigonometry were provided by the Special Education section of the Austin Public School System. Other equipment included water colors, crayons, white and colored chalk, construction paper, tablets, typing paper, molding clay, pencils, finger paints, magazines and scrapbook making material.

LIMITATIONS

All too often a mental institution is a catchall for everything society deems undesirable enough to segregate from itself. In order to make a hospital an instrument of rehabilitation some degree of discrimination must be maintained. If it is necessary, because of inadequacies in the legal system, for the hospital to cater to the subnormals as well as the abnormals, the hospital must make a frank differentiation between the programs to be followed in treatment of these two categories. The purpose of Educational Therapy programs is to restore the patient at least to his former level of efficiency, and, if possible, to increase his level of functioning; but it must be recognized that no educational technique invented has ever replaced a missing gray cell. By trying to make one program flexible enough to include every form of deviation the net result will be reduced to zero.

This is not intended to relegate the subnormals to the wastebasket. It merely segregates them from the group under discussion. The subnormals form a group which is custodial in nature, and their inclusion in a dynamic program for readmission to normal society will lead only to a blind alley.

The idea of having the school open to voluntary attendance by any and all the hospital patients was advanced but seemed impractical on at least six counts. (1) The nature of the students makes it imperative that the classes be small so that individual attention can be given to the students. On several occasions this has been borne out substantially when it was found that two students could not be given simultaneous instruction on parallel

assignments. This is not strange in that, perhaps, those who became students in this school were those who were unable to function adequately in their society, thus it is small wonder that they find group participation difficult—even on this limited scale.

(2) Another difficulty which is closely related to the above is the definite amount of work which an assignment can contain for each individual. If any assignment exceeds this amount it is rejected as being too difficult, and the student feels 'put upon'. If the assignment contains less than the optimum it is within easy reach of the student which creates for the teacher the problem of keeping everyone occupied. With a large class the teacher cannot get around to all before they have finished such assignments.

(3) These students are more insistent than the normal students about having their work checked and graded immediately. They lack courtesy and tact and will interrupt any activity in progress to secure their most recent desire. A large class of this type would soon deteriorate into a condition of bedlam.

(4) Another practical disadvantage to having 'open' class is the disquieting effect (to say the least) created when the teacher walks into a room filled with students about whom he knows nothing except that they must be abnormal or they wouldn't be there. Not knowing their educational achievements or their idiosyncrasies the teacher is at a loss to know how and where to start and the direction to take.

If the teacher starts a paranoid too low his pride is injured and rapport is extremely difficult to establish subsequently. The same amount of coercion-cajoling that can be used successfully with a schizoid meets with disaster if used on a hostile organic student.

A positive program of admission to the school which was worked out and adhered to circumvented many of these unfortunate situations. Opportunity for the teacher to brief himself on each case before the student was admitted to the classroom saved confusion and embarrassment in all directions.

(5) Equipment, books especially, are taken to the wards at will because it is felt that this permission gives the students a sense of belonging and being accepted. The free flow of students, however, makes this arrangement obviously inadvisable.

(6) Several of the students are interested in making the class a meeting place for themselves and their *objets d' amour*, and little

or no academic ambition enters the picture. They easily become a disrupting influence for, if they are put to work, they take the limited facilities from those who have genuine academic aspirations, and if they are not put to work, they undermine the entire tenor of the class with their erotic activities.

Thus it was felt to be in the best interests to limit the students in the class, and to this effect a screening criterion was set up. It might be noted that as things turned out this standard was frequently honored in the breach. Four categories of patients were set aside as being inadvisable: (1) those of advanced age. These were sufficiently removed from school by virtue of their years that it was felt their inclusion in school would tend to be a regressing factor rather than a normalizing factor. In the case of the younger patients, even though their academic progress might be slight by comparison to outside standards, it was felt that for them to engage in an activity which resembled, at least to some degree, their concept of normal outside work would be in some measure therapeutic.

(2) Those currently engaged in a course of shock treatment were eliminated also. The advisability of this was subsequently underlined by the exception. One girl who was engaged in shock therapy disrupted classes thoroughly; it was impossible to arrest her attention even for short intervals, she could remember none of the names of the classmates (even though her avowed true love was among them) and in general she macerated academic progress.

(3) Those with organic involvement (brain damage), especially of the progressive types, were placed on the nonpreferred list because it was thought that their inclusion with others of greater capacity would be beyond the point of diminishing returns with respect to their administration. There were three, however, who were admitted to the class: one trauma and two lobotomies. The trauma case did little of academic interest and was much more concerned with sex relations than arithmetic. One lobotomy made great strides toward regaining proficiency in geometry, but the other was on shock treatment also and the summative effect was that she was considerably beyond the pale of advisability for our class.

(4) The psychopathic personalities were frowned upon mostly because of their anti-social activities. When all is said and done,

it is still necessary to maintain some sort of order in the classroom. To maximize the effectiveness of the program, it was thought advisable to minimize the number of psychopaths.

AIMS AND PURPOSES

Education as therapy upholds both the principles of modern education and the tenets of good mental health; for both are based on the same concept—capitalization on individual differences—but neither science escapes modification by the union. Our purpose is not purely education. If it were this experimental project would be unnecessary; any typical classroom situation would suffice. Our students were frankly abnormal—beyond the limits of tolerance in normal classrooms; so it seems advisable to define the aims and purposes which guided the experimental class.

1) To provide basic academic skills for the student to use to facilitate his future adjustment.

For anyone to succeed in this day as even a somewhat autonomous creature, he must have some control over his environment. It is held to be almost patent that one who has attained a chronological development far in excess of one's academic progress is at a distinct disadvantage in one's dealings with inanimate objects and especially with interpersonal relations. If there is to be any degree of rehabilitation for these young students, it will necessarily follow that they compensate in some measure for their scholastic deficiency incurred as a result of their mental illness.

The likelihood of their acceptance into and their employment by society progressively declines with retardation in academic progress. Who, in his right mind, would employ a patient for any sort of desirable position if the patient could not read or write or tell time? If the patient is destined for nothing more than custodial care for life, these things become of trifling importance for any environment which is equipped to care for him is equipped to make allowances for his inabilities. If he is to become a responsible citizen, he must meet the basic standards.

Intellectual performance by an individual suffering from a mental condition is almost always impaired, and the encroachment is usually spotty and uneven. The students themselves are unaware of their deficiencies; so it must be up to the teacher to ascertain their lacks (and some of these show up only after considerable time)

and to fill the gaps so that additional increment will be based on a solid structure of accomplishment.

2) To provide basic personal skills for the student to use to facilitate his future adjustment.

It seems that one of the first things to go when an individual is besieged by mental illness is his will to make himself a part of cleanliness and neatness. He lets his clothes become dirty, disorderly and ragged; his hair and teeth go unmolested by comb and brush; his school work and letter-writing go to seed; and water seems of little concern to him since a great portion of that which gets on his outside is only by accident.

Since we became a germ-conscious culture some years ago, there has been a compulsion toward neatness which distinguishes our middle class from almost every other subculture on earth. From the point of view of the world outside the hospital a degree of personal neatness not commonly found among hospital populations will be a prerequisite to satisfactory adjustment.

It is recognized that a superficial veneer of tidiness will do little toward amelioration of the basic disorder, but it is felt that gentle urging in that direction, setting examples for the students and the constant recognition of neatness as a symptom of normalcy will do a great deal toward reinstating the student's self-respect when other more pressing problems are resolved.

There is to be considered also the hygienic value of cleanliness and the concomitant morbid mental state that frequently goes with, or is aggravated by, a microgenic disorder; but these are so commonly accepted that they need only be mentioned.

3) To provide accredited academic work on primary and secondary levels.

Our class was fortunate in having the coöperation of the Austin Public School System in the accrediting of work done in school; for credit, *per se*, is important to the student as well as to others. For the student it provides some tangible goal for him to strive toward; some milestone which he can use to signify a definite level of achievement to satisfy himself that he is not merely vegetating during his stay in the hospital.

Most establishments which would be prospective employers of the students have no program for testing the proficiency of the

students along these lines; so the symbol of proficiency, the diploma, frequently becomes the universal, critical key to the future.

4) To provide information which will serve as a basis for their discrimination in introspection.

It has been noticed that some of the students realize that they have had hallucinations, and the only basis for discrimination between facts and artifacts is their fund of information. The realization that tinnitus is a universal occurrence and not an indication of malady came as a startling revelation to one student. Such understanding of natural phenomenon and the significance of introspective observations will go a long way toward minimizing misinterpretation and should increase the student's grip on reality as he begins to gain some insight into himself.

5) To occupy the student constructively.

It has been said, "No wonder epileptics deteriorate: they have nothing to do but sit around all day and watch each other throw fits." In the same vein, residents of mental hospitals who have nothing to do but contemplate their symptoms and vegetate in the behavioral fertilizer of other psychotics can hardly be expected to derive much benefit from their contacts. Educational therapy is designed, at best, to provide goals and incentives to reach those goals—at least, to give the students something to keep them from perseverating on their own morbidity.

6) To provide activities and entertainment which are designed to keep the student in contact with his concept of normalcy.

It has been noted that some of the students who gain a bit of insight into their conditions identify themselves with the abnormal. For these students to reach a high degree of normal functioning while isolated in the hospital requires a bootstrap sort of effort. They must identify with normalcy without being exposed to it. The process of deterioration may be nothing more than an internalization of the lax and permissive attitudes of the institution. Abandoning hope for release and realizing that he need have no rigid standards to continue life in the hospital, the student gives up the struggle.

Educational therapy is designed to bridge the gap between normalcy with its intolerance and hospital permissiveness. It

brings the 'outside', non-threateningly, to the student for his acceptance.

7) To provide opportunities for social interaction.

In addition to identifying personally with normalcy an important goal is that of identifying with society as a collection of individuals. The pseudotolerance which is often exhibited by mental patients toward the behavioral peculiarities of others is not based on acceptance of them, but rather on a wide separation of their respective inner worlds. When these separate worlds of action are brought into juxtaposition, it is usually observed that they react in a characteristic short-sighted, egocentric manner.

If the students could be brought to some degree of realization that one can make allowances for others without loss of status all would be well; but if this is impossible, it might still serve a valuable purpose if one were taught merely to inhibit one's aggressions, thus obviating the vicious cycle of threat and ego defense. It was found possible to use the rapport which was established to circumvent some threatening unpleasantness. It is felt that as *esprit de corps* develops that social pressures will come to mean more in shaping their activities.

In order to develop this feeling, activities which set the students apart from the rest of the hospital population were planned. They included movies on a wide variety of subjects, student-selected field trips such as the museum and Barton Springs (one student caused a furor when he announced to the life guard that he was nine feet tall, weighed three hundred and ten pounds and was the best looking man on the grounds).

There is also in class an opportunity for students of opposite sexes to meet in supervised but permissive atmosphere. Such constitutes experience which should be in the repertoire of the student before he reënters society.

8) To provide opportunities for success.

It is so important that the students be made to feel successful in their activities that rather than being merely a technique for achieving other aims, this is in itself a goal. There is an effort made not to assign any task which is beyond the realm of attainment for the student. He is given no deadlines for its completion and much commendation is rendered when it is finished. If it is only

partly correct, it is considered correct with an explanation of the alterations necessary. If it is incorrect in its entirety, no blame or censure or disapproval is indicated, especially when the student seems to be chastising himself, and encouragement and reassurance is always abundant.

Grades are optional with the student. Some prefer them and others do not, but in all cases a pittance of work nets a superlative grade. Seldom are they less than 95, and when work is corrected only the second mark, 100 is given.

It has been noticed that these students frequently prefer dull, routine drill. This seems to be a striving for adequacy and security, and it must be catered to before the desire on the part of others for a greater academic achievement. It is hoped that when they feel a genuine security the students will show spontaneous tendencies toward exploration of unknown fields and diversification.

9) To provide an accepting atmosphere.

An accepting atmosphere is indispensable to rapport, to a feeling of non-threat, and hence to a facile learning situation. The key to this seems to be in quiet, interested, unemotional receiving of any novel and startling innovation in student behavior. All the eyebrow-raising details are greeted without panic, and they are referred to thereafter only by the students. Quite naturally, in this population there is a limit beyond which aggression and negativism cannot be tolerated, but, happily, these limits are seldom reached.

On 'those days' students are allowed to read comics, just sit and stare, return to the wards, sleep under the tables or whatever they feel like doing. They are asked frequently if they have suggestions and practically any spontaneous activity is included in the curriculum.

Competition, whether it be between the students or between school and students, is not sponsored. Each student works on his particular assignment and is paced by his own ability rather than any frantic sense of academic inferiority-avoidance which is so frequently used to motivate normal students.

To increase the atmosphere of the school and diminish the hospital influence the students are referred to as such and never as patients.

At first the students seemed to get nowhere and some used this

permissiveness as a loafing device throughout, but most seemed to do as well as their limitations would allow.

10) To collect data which may shed light on these conditions, their amelioration, education and perhaps their prevention.

In all the literature on the education of exceptional individuals there seems to be a paucity concerning mental disorders, perhaps due to lack of contact with truly abnormal people in the average educational environment. Hypotheses which have been made and on which therapy has been based can be evaluated only by comparisons yielded from this and comparable classes; so for this our class supplied a daily log of events and attitudes, a collection of their work and other miscellaneous items: deposits for the future.

RESULTS

The class met on one hundred days from April through August, and had enrolled eleven students whose attendance ranged from four to ninety-six days. One boy who attended class only one day is not counted.

Academic progress was generally better than had been anticipated. The most outstanding advance was made in typing. The student progressed from an incomplete knowledge of the keyboard to long paragraphs of mistake-free transcription done rhythmically and with promise of speed.

Others showed considerably less improvement. One had hope of mastering the Arabic numeral system from 1 to 12 preparatory to learning to tell time. He managed nine numbers in only one hundred days which was in reality an accomplishment.

Five students made notable progress, either in new subjects or in recapturing departed abilities. One who showed great promise left the warm and mothering confines after only four days of attendance.

Five others showed somewhat less progress, two of these were the lobotomies, one was on shock treatment, and one was the brain damage case. The two aments failed to grow in academic grace, but they did not shrink either.

The personality progress of these students is harder to ascertain. An effective appraisal would necessitate frequent observations by one trained for that purpose, but lacking these facilities, the only alternative was to base the progress reports on the daily contacts

in the school. Five made strides toward gaining self-confidence, socialization and development of self-control. The two who were aments were never a personality problem. The organic case was mostly uninfluenced—as might be supposed. Three showed no improvement, but, happily, no deterioration either.

FUTURE CLASSES

One group omitted from the experimental program would seem to merit special attention. Here in the South Negroes generally maintain an average intellectual achievement lower than that of whites. It would often be an unrewarding task to try to bring the educational average of the Negroes up to the level of the whites, for if a Negro were returned to his peer group with an educational level which was very far in excess of the group he probably would find acceptance by even his own group more difficult—and acceptance by the whites still impossible—thus aggravating the very problem which is crying for solution.

Luckily, the absolute level of education is of much less importance than the relative level. If he returned with a new ability—perhaps the ability to read where before he was illiterate—he would be in a position to command increased respect from his associates and still be accepted.

It would seem clear that the ultimate goals of each case would be based on a thorough study of the case itself and how it articulates with its environment, rather than a remote theoretical criterion.

LIFE ACTIVITIES

The feeling of accomplishment for anyone might also go a long way toward counteracting the stigma of being a mental patient which is all too often still a part of present-day thinking. The patient who is forced to admit, even though only to himself, that he is handicapped in his living because of losses due to his mental illness, is placed at a distinct disadvantage; but if he could say, "Look! I can perform better because of my stay in the hospital," he would be better able to face the readjustment to society and the acceptance of his own illness.

Many things which the non-institutionalized person takes for granted could be the source of acute embarrassment to one who has been out of contact with society for some time. For example,

what embarrassment succeeds a frustrating encounter with the ubiquitous coin machine which even the little children can operate? Likewise the cataclysms which could be caused by the eye operated doors, help-yourself supermarkets, late model automobiles, multi-directional traffic signals, express busses, self-operated elevators, perhaps even dial telephones and the other myriads of gadgets which characterize our age.

For one engaged in the herculean task of readjusting to a life that has passed him by, many such problems might be critical in causing existence either to be smooth and flowing or a succession of accute discomforts.

Toward the end of his stay classes would include 'Living' which stresses smooth manipulation of common situations. The intensity of this phase would depend on the total time he has been effectively removed from society, the state he had reached in his premorbid social development and social expectancy.

ORGANIZATION

The inclusion of Educational Therapy would not eliminate such therapies as Occupational, Recreational and Music, nor would it merely superimpose another one of these on the existing structure. Rather it would integrate them into a single, coördinated, multi-directional therapy for a concentrated attack on the illness. Educational Therapy should include all these and anything else that would tend to increase the individual's ability to perform successfully in life.

Because of the variety of backgrounds and the differences in goals of those patients admitted to the Educational program it cannot take on a linear form such as is found in ordinary schools. These assume that everyone begins at the same level and ends at the same level. Here, however, an areal pattern would obtain which would allow one to start at his own level and work upward through his stay in the hospital toward individual goals.

Upon entry into the hospital the usual legal, medical and psychological processing would take place as rapidly as possible. Part of this processing would include information concerning his academic past to assist in his placement. Also included would be items which would shed light on his premorbid educational expectations to be used as aids in fixing his goals, and they would be influenced

also by his mental capacity, his degree of adjustment, rate of improvement and social expectancy.

His first classes would be simple, small, informal, permissive and homogenous groups (perhaps cleanliness only to accompany shock treatment). His success would be the desideratum. Classes should include as much variety as possible and still be within easy reach of the student. Their periods should be short, not more than thirty minutes, to accommodate their restricted attention spans. Periodically some physical activity must be provided to assuage the lethargy in the withdrawn ones and to give release to the hyperactive ones; also it has been observed that movies of almost any kind have a general appeal, almost magical, and good results have been experienced with having refreshments (cookies, pop, graham crackers, fruit juices, etc.) at odd times during the day. One half day of classes should be considered a full load at the beginning and as a student's disturbance abates there would be a corresponding increase in the academic load. Students in the experimental class responded well to the idea of having free access to the equipment and being placed on their honor as to its use.

Although permissiveness is the key to the first courses, it must go hand-in-hand with persuasion and a great deal of encouragement. It must be remembered that, for the patient, this is not his major battle; so he may be expected to lack a great deal of enthusiasm for school. These first classes must be limited to six or eight, and even fewer if they are severely disturbed or highly maniac.

As the student progresses and is better able to make successful adjustment to additional responsibilities, he will be promoted to a larger class with more exacting standards. This graded permissivness insures the warmth necessary to the beginner, and gradually 'toughens' him so that he will not be overwhelmed upon release from the hospital.

Concurrently, of course, the psycho-environmental and medical aspects would be operative while treatment of his milieu would be progressing behind the scene. Medicine has long believed that medication followed by reinfection was effort wasted. Similarly here. The social services would have as one of it's chief functions either reëducation of the patient's former environment to a point where it no longer constituted a threat to him, or placing the student where he can be accepted. Graduation would be synonymous

with discharge from the hospital, and a diploma should indicate work accomplished.

TEACHER

The teachers for this program will constitute the major problem. If the endeavor is ever to be successful, it must be undertaken completely—which includes an adequate student/therapist ratio. The teachers for such a program would be well-grounded in psychology and the basic concepts of psychotherapy. When all is said and done it is still an organization devoted to the restoration of mental health; education here is merely the key to this restoration, and nothing should obscure this salient fact. The teacher should be acquainted with educational procedures, especially those which deal with the education of exceptional persons.

These two factors alone are not enough. Especially for the first classes, the therapist must have a warm, easy-going, 'mushy' personality which will allow him to identify easily with his students and they with him. Students respond much better to a bit of horizontal interest than to a business like, rigorous, discipline-based authority. They seem to accomplish more in less time when allowed to spend some time in ventilating their feelings in class.

SUMMARY

This is a report on an experimental class in Education Therapy which was conducted at the Austin State Hospital in 1951. The goals of the program were (1) to shorten the stay in the hospital of all who are remissable, (2) to provide a stable viaduct back to normalcy facilitating readjustment, (3) thereby reducing the number of readmissions and (4) eventually reducing the over-all number of mental patients in hospitals.

The chief tool in the task of preparation and readjustment would be an integrated educational program. The aims and purposes of this academic situation in particular, and perhaps for Educational Therapy in general are:

1) To provide basic academic skills for the student to use to facilitate his future adjustment.

2) To provide basic personal skills for the student to use to facilitate his future adjustment.

3) To provide accredited academic work for primary and secondary levels.

4) To provide information which will serve as a basis for their discrimination in introspection.

5) To occupy the students constructively.

6) To provide activities and entertainment which are designed to keep the student in contact with his concept of normalcy.

7) To provide opportunities for social interaction.

8) To provide opportunities for success.

9) To provide an accepting atmosphere.

10) To collect data which may shed light on these conditions, their amelioration, education and perhaps their prevention.

WHO IS PENALIZED BY THE PENALTY FOR GUESSING?

ALEX C. SHERRIFFS and DONALD S. BOOMER

University of California

Almost since the introduction of objective tests into the field of educational measurement, teachers have been concerned with the guessing factor as it enters into test scores. Correction formulae of varying degrees of complexity have been developed. All are variations of the original method of subtracting the number wrong from the number right, ignoring omitted items, and considering the remainder as the significant measure.

Generally it has been shown that the application of correction formulae results in a wider range of test scores and in downward adjustment of average scores. Some have claimed that chance effects are reduced and that a more reliable set of scores results (*5*). Others, however, believe that a correction formula lowers reliability (*4*, *6*). Ruch (*7*), after surveying the literature to 1929, concludes that the effect of a correction term on reliability is debatable.

All in all, though the effects of the theoretical corrections for chance are completely unclear, such corrections are applied by many teachers in the classroom situation. One extreme defender of the penalty for guessing aims his correction at the 'gamblers', 'bluffers', and 'bullies' among the students, stating that "The results (of penalizing wrong answers) are always the same. Students gradually acquire a new regard for truth, precision of learning, the impropriety of guessing, the shame of bluffing, *et cetera*" (*3*).

The present writers have for some time believed that the usual notions about the effects of right-minus-wrong scoring have ignored some important factors. Our own reasoning has been somewhat as follows.

A course examination is administered in order to evaluate the present 'achievement' of students—their knowledge of course content, or their ability to utilize their course experience in relevant problem situations.

It is our contention that course examinations should not serve as direct tests of intelligence. (If measurement of intelligence is

desired, quantities of standardized tests for the purpose are available.) Neither ought they be measures of the student's personality, nor should successful performance in the testing situation be contaminated any more than necessary by personality variables.[1]

We believe that when the right-minus-wrong instructions are added to an ordinary examination a new set of decisions faces the individual student. For example, such questions must necessarily arise as: "How certain am I of this answer? What is the probability of my being right? If the probability is 80 to 20 that I am right, then am I better off to 'guess' or to omit the item?" Ability to 'figure the odds', subjective levels of confidence in own judgment, and willingness to gamble or to take a chance, are among the variables which might well be expected to influence the right-minus-wrong measure of achievement. Such influence can only make the test less valid as a measure of achievement.

It seems obvious that such a contaminated measure is not only less useful for its avowed purpose, but also that it will work hardships on individual students because of individual differences in shrewdness in test-taking or in other personal factors.

The present study represents an attempt to find evidence as to the presence of contaminating factors in the conventional right-minus-wrong test score.

PROCEDURE

Two approaches to the question under consideration suggest themselves. One approach involves scoring the performance of each student on one test in two ways—by applying a penalty for guessing, and by a simple count of the number of items answered correctly. The two resulting sets of scores can then be compared in their relation to each other and to other pertinent variables. A second approach would involve having some course examinations of the right-minus-wrong type and others of the ordinary 'no

[1] We agree with Cronbach (*1*) that "When the test is used to fill other purposes, the score becomes invalid for the primary purpose of measurement; essentially the more elements that enter a score, the less meaningful it becomes." A few writers have considered the dangers of contaminating measures of course achievement with the effects of personality differences (8, *10*, *13*) but only Votaw (*11*) has approached this question experimentally. He presents data on the relationship of measures of dominance-submission to guessing behavior.

penalty' variety; then comparing the student's scores on the two types of tests.

There are advantages and disadvantages inherent in both designs. The first approach was followed in this investigation.

Subjects.—The subjects were students in the second semester course in psychology designed for non-psychology majors at the University of California. Of the four hundred and fifty students who were enrolled in this course three hundred and ninety are active subjects in this investigation. The criteria for inclusion were that the students have participated in four midterms and the final examination and that the students have satisfactorily completed the Minnesota Multiphasic Psychological Inventory.

Experimental Examination—The first midterm examination consisting of fifty true-false items served as the experimental examination for this study. The class was divided between two examination rooms, students whose last names fell in the first five letters of the alphabet going to one room, the remainder to the other.

In one room ninety-eight subjects took the examination on a 'no penalty for guessing' basis. "Answer all questions, please." This group served as a control for the test-taking condition.

In the other room two hundred and ninety-two subjects were told that the examination was to be scored on a right-minus-wrong basis with no score for omitted items. After these students had completed their tests they were instructed to circle all the items they had omitted. Following the completion of this task, they were asked now to indicate what they believed to be the most probable right answers for the previously omitted items. The circling technique made it possible to obtain three scores for each student's performance on the examination. These were (1) a right-minus-wrong score (R−W), (2) a score representing the total number of items that the student answered correctly when required to try all (No. Right), and (3) a penalty score (Penalty Score) determined for each student by the difference between his R−W score and his No. Right score.

Criteria for course achievement.—To aid in interpretation of results some definition of course achievement was desired. A best approximation was utilized for the study, consisting of the combined scores on the three midterms following the experimental examination and the final examination. True-false questions,

multiple-choice questions and essay questions each contributed about one-third to this measure. This measure of course achievement turned out to be stable, to yield a wide distribution of scores for our sample of subjects, and, in terms of subjective evaluations of staff, generally to rank students in accord with expectations based upon office hour contacts.

Personality dimension.—It was our hypothesis that students who lacked self-confidence, especially confidence in their own judgment, who had difficulties in making decisions, and who were easily threatened by ambiguous situations, would be penalized by the R − W test condition. To test this hypothesis we desired a personality measure which might be expected to relate to some of these variables.

The measure determined upon was the A scale of the MMPI. This is a new scale of thirty-nine items developed by Welsh (*12*). It is claimed that people scoring high on this scale are characterized by introversion, rumination, anxiety, low self-esteem, and undue concern with the impression they make on others. The 'flavor' of the scale, its face-validity, is conveyed by such items as:

> Criticism or scolding hurts me terribly (true)
> I worry quite a bit over possible misfortunes (true)
> I have often lost out on things because I couldn't make up my mind soon enough (true)

The group form of the MMPI was administered to the entire class in mid-semester, and for the purpose of this investigation was scored for the 'A scale'.[2]

RESULTS

Experimental Groups

1) *'R − W' and 'No Right' as measures of achievement.*—Comparisons were made of the two scoring methods applied to the fifty-point experimental midterm. The Pearson Product Moment correlation between the 'R − W' and 'No Rt' measures was .96.

[2] The investigators did not believe, of course, that the 'A scale' was the only measure worth following up, nor did they believe that they knew at the present stage of its validation exactly what aspects of personality it measured. However, available evidence about this scale suggested it to be a good bet. Insofar as the case against 'R-W' scoring goes, contamination by any personality factors, whether or not closely defined, would be important.

The correlation between the 'R − W' scoring of the experimental midterm and the criterion measure of course achievement was .48; between 'No Rt' and the criterion the correlation was .49.

The 'R − W' scoring method yielded a lower mean score and a more variable distribution of scores. However, this must be seen in relation to the extremely high correlation between 'R − W' and 'No Rt' scoring, and essentially the same degree of relationship between these scoring methods and the criterion for course achievement.

2) *The method of scoring and the personality measure.*—The 'A scale' yielded a skewed distribution of scores with a mean of 9.93, a standard deviation of 7.8, and a range from 0 to 35.

Correlation coefficients were calculated in an effort to assess the degree of relationship between the 'A scale' score and the two scoring approaches to the experimental midterm. The correlation between 'A' and 'R − W' was −.23; between 'A' and 'No Rt' it was −.22. A coefficient was also derived for the relationship between the 'A scale' and the achievement criterion. This correlation was −.20. These correlations are all significantly different from zero.[3]

The above findings do not answer the question as to whether individual students were differentially handicapped by 'R − W' scoring because of personality factors. A comparison of those students high in maladjustment as indicated by the 'A scale' with those low should, however, offer crucial evidence. It was decided to study approximately fifty cases at each end of the distribution. Because of the presence of a relationship between the personality variable and course and examination performance, subjects from the two ends of the distribution were paired on the basis of their total course achievement. It was possible to obtain forty-seven pairs, matched within two points on the achievement criterion. Thus we are comparing two groups of students who functioned equally well on the accepted measure of course mastery and who differ only in that they are the 'highs' and the 'lows' on the personality variable in question. How these groups functioned on the experimental midterm is summarized in Table I.

---

[3] Examination of the scatter plots reveals the relationship between the 'A scale' and all other measures to be curvilinear. Thus the reported correlations can only be used to predict in one direction—from the 'A scale' to the performance measures.

TABLE I.—COMPARISON OF PERFORMANCE ON EXPERIMENTAL MIDTERM OF STUDENTS WITH 'A SCALE' SCORES OF 15 AND OVER WITH THOSE WITH SCORES OF 4 AND LESS

	Highs (A scale score of 15 and greater)	*Lows* (A scale score of 4 and less)	Mean Diff.	Std. Error	CR
Mean R-W Score	18.87	22.85	3.98	1.23	3.23
Mean No. Rt. Score	34.89	36.32	1.43	.89	1.61
Mean *Penalty* Score*	16.02	13.47	2.55	.837	3.05

* The *penalty* score represents the difference between the gross scores on R-W and on No. Rt. It is not a completely satisfactory statistic. However, the skewness of the R-W distribution made questionable the use of differences between standard scores.

These findings show that students with high 'A' scores were penalized by the use of 'R − W' scoring in comparison with students whose 'A scale' scores were low. The difference between these groups is statistically significant at better than the .01 level. Further, this is a difference of practical significance, amounting to almost a full letter grade within the 'R − W' distribution.

3) *Analysis of patterns of performance under the R − W condition.*—A study of the 'omitting' behavior of the subjects of the experimental group was made in an effort to throw additional light on the manner in which the higher 'A scale' student was penalized.

One hundred and eighty-seven students omitted three or more items. These subjects were divided into two groups: those whose 'A scale' scores were 15 or over and those with scores of 14 and less. Next, these subjects were paired, high 'A' with low 'A' on the basis of their 'No Rt' scores. This was to match them in terms of their knowledge of the material covered by the examination. It was possible to obtain forty-four perfectly matched pairs.

Two comparisons were then made. The first was in terms of the number of items omitted by the two groups. Students with higher 'A scale' scores omitted under the 'R − W' conditions 6.25 items, on the average; those with low 'A scale' scores omitted 5.09. This difference results in a critical ratio of 2.03 and is significant at the .05 level.

A second comparison was made. The proportion of items later answered correctly (after circling) by the two groups was studied

to see whether the higher 'A' subjects differed from the low 'A' subjects in this regard. Sixty-three per cent of the items omitted by the higher 'A' subjects were later answered correctly by them. Fifty per cent of omitted items were later answered correctly by the low 'A' subjects. The significance of this difference of thirteen per cent was assessed in three ways: In terms of the *t* resulting from analysis of correlated proportions, in terms of the difference between paired percentages utilizing the $\sqrt{\text{arc-sine}}$ transformation, and according to the non-parametric 'Sign Test' suggested by Dixon and Mood (*2*). In all three cases this difference proved significant at the .01 level or better.

These findings reveal that the excess penalty which we have found to accrue to more maladjusted students comes in part, at least, from two factors. The more maladjusted students omit more items, and they omit a larger proportion of items that they are capable of answering.[4]

Control Group

It will be recalled that ninety-six of our subjects took the identical midterm as did the experimental group, but under different conditions. The controls never believed the examination to be of the penalty-for-guessing type, taking it only on the 'try all the questions' basis.

The control group data provides information making possible sharper interpretation of certain of the experimental group results.

1) Did the experimental group subjects try less hard on the previously omitted (circled) items when instructed to answer them than they would have done had the examination been a conventional 'number right' type in the first place?

2) Did the very situation of being confronted with a 'R − W' test modify the performances of students in general?

---

[4] Additional information was obtained on this matter by taking another slice at the data. Sixty-six pairs of students were studied. These were matched for 'No. Rt.', but divided into two groups: those who answered correctly better than fifty per cent of those they omitted, and those who answered correctly only fifty per cent or less. A comparison of the 'A scale' scores for those two groups revealed a mean 'A scale' score for those students more successful on omitted items of 11.73, and a mean score for those less successful of 8.42. The mean difference of 3.31 points has a standard error of 1.25 points. The Critical Ratio for this difference is 2.65. The difference, then, is significant at the .01 level.

TABLE 2.—COMPARISON OF CONTROL AND EXPERIMENTAL GROUPS ON BASIC VARIABLES

	Achievement Criterion		No. Rt. on Experimental Exam		A scale Scores	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Control subjects (n = 96)	242.97	20.61	35.50	4.40	10.07	8.51
Experimental subjects (n = 292)	241.87	19.59	35.87	4.21	9.95	7.80

3) In the case of those students who had been found to be handicapped by the penalty-for-guessing scoring, was part of the explanation of this handicap to be found in the general atmosphere of the test situation itself?

Should the control group turn in a better performance on the examination than did the experimental group, the answer to any or all of the above questions could be 'yes'. It would be difficult with our data to determine the relative contributions of carelessness and of insecurity. However, should there prove to be no difference, then the indications would be negative on both counts.

A comparison of the control and experimental groups follows.

None of these differences between the Experimental and Control groups was significant at better than the .30 level.

In the case of the control group a correlation of .56 was found between the experimental midterm and the criterion for achievement. When this is compared with the .49 in the case of the experimental group, the difference is not significant.[5]

Two final checks were made. Those subjects of the control group who had 'A scale' scores of 15 and over were compared in their performance with those who had scores of 14 and less. Ad-

[5] The correlation between the 'A scale' and the control group scores on the experimental midterm was —.003. Between the 'A scale' and the criterion for achievement it was —.070.

The significance of the difference between these 'zero' control correlations and the comparable experimental ones (of the order of —.20) was calculated utilizing Fisher's Transformation. These differences do reach the .05 level of significance. We are unable to find an explanation for these differences.

ditionally, control subjects were matched for 'A scale' score with cases from the experimental group. All other variables were disregarded during the pairing. No significant difference resulted from any comparison. High 'A' subjects in the control group were not different from low 'A' subjects in regard to 'No Rt' on the experimental midterm, or in regard to course achievement. High 'A' subjects in the control group were not different from high 'A' subjects in the experimental group, nor were low 'A' controls different from low 'A' experimental subjects.

Study of the control group, then, leads us to the conclusion that there is no evidence of carelessness, or of disturbance in 'No Rt' performance simply from being in the 'R − W' testing situation.

CONCLUSIONS

On the basis of this investigation we find no reason to reject our original hypothesis. We find that students who score on the maladjusted end of the 'A scale' are handicapped by being scored on a 'R − W' basis. This holds true even when the students' knowledge of course material is held constant. The evidence suggests that these students are penalized by their tendency to omit more items, and to omit items the answers to which they know.

We can see no justification for teachers continuing to use corrections for guessing. The assets accruing to such a system have always been equivocal, and the liabilities, we believe, are now apparent.

We hope that Traxler (*9*) is unduly pessimistic when he says that he doubts that school people will ever be willing to encourage pupils to guess at every item in a test. He believes that these educators would feel this to be condoning carelessness and loose thinking. We hope that the educator will also take into account principles of fairness, and of respect for the 'whole child'.

REFERENCES

1) L. J. Cronbach, "True-false tests: A reply to Count Etoxinod." *Education*, 1941, 62, 59–61.

2) W. J. Dixon and A. M. Mood, "The statistical sign test." *J. Amer. Stat. Assn.*, 1946, 41, 557–566.

3) S. Etoxinod, "How to checkmate certain vicious consequences of true-false tests." *Education*, 1940, 61, 223–227.

4) Harry A. Greene and A. N. Jorgensen, *The Use and Interpretation of High-School Tests*. New York: Longmans, Green & Co., 1935.

5) Herbert E. Hawkes, E. F. Linquist and C. R. Mann. *The Construction and Use of Achievement Examinations*. Boston: Houghton-Mifflin Co., 1936.

6) C. C. Moore, "Rights minus wrongs method of correcting chance factors in the true-false examination." *Pedagogical Seminary*, 1940, 57, 317–326.

7) G. M. Ruch, *The Objective or New Type-Examination*. Chicago: Scott, Foresman & Co., 1929.

8) Frances Swineford, "The measurement of a personality trait." *J. Educ. Psychol.*, 1938, 29, 295–300.

9) J. W. Varty, "Guessing on examinations—Is it worth while?" *Educ. Forum*, 1946, 10, 205–212.

10) D. F. Votaw. "The effect of do-not-guess directions on the validity of true-false or multiple-choice tests." *J. Educ. Psychol.*, 1936, 698–703.

11) George S. Welsh. Personal communication to writers.

12) Llewellyn N. Wiley and Otis C. Trimble, "The ordinary objective test as a possible criterion of certain personality traits." *School and Society*, 1936, 43, 446–448.

TOPOGRAPHICAL ORIENTATION IN NAVAL AVIATION CADETS[1]

BRANT CLARK AND R. DANIEL MALONE

United States Naval School of Aviation Medicine, Pensacola, Florida

The problem of how an individual orients himself to geographical points is an old one and has involved much theoretical controversy. Nevertheless, very little clear understanding of the specific processes involved has resulted, although radical postulations of 'special senses' in both animals and man have been common. The most intensive studies have been made by educators who have been interested in geography (*4, 15, 16, 18, 19, 23, 27*). The majority of these studies have investigated childrens' ability to point north or to other compass points, while some of them have been concerned with the childrens' ability to indicate the direction of places within their own community as well as distant points. Only a half dozen studies have been found which were concerned with topographical orientation in human adults. Two of these papers were essentially anecdotal in nature and merely describe the activities of a limited number of individuals having either a high degree of spatial orientation or experiences of being seriously disoriented (*2, 3*). Only two investigations (*14, 31*) have employed a large number of subjects. These studies have demonstrated wide individual differences in orientation to geographical points as well as constant group errors. It has been demonstrated that scores on tests of orientation to geographical points show low correlations with general intelligence, age, knowledge of the vicinity, sex and school experience. Vision and mental maps appear to play a predominate rôle in aiding the individual in orientation.

The importance of adequate orientation is well-known to pilots of aircraft (*7, 8, 22, 26, 28*). This has been demonstrated by analyses of flight safety records, (*7, 26, 28*) analyses of the causes of failures in flight training, (*17*) and the relationship between spatial and visual orientation to success in flight training (*17, 20*). Experiences reported by pilots (*28*) show that in spite of their very

[1] Opinions or conclusions in this report are those of the authors. They are not to be construed as necessarily reflecting the view or the endorsement of the Navy Department.

efficient navigational aids, pilots do occasionally become confused with regard to their geographical position. It is significant to note that when this disorientation occurs, it is associated with marked nervousness and uncertainty on the part of the individual (*2*, *28*).

It was the purpose of this study: 1) to determine how well Naval Aviation cadets are oriented to geographical points; 2) to identify common errors in orientation; and 3) to compare topographical orientation with two other tests of spatial orientation, tests of other mental abilities, and certain selected personal data.

PROCEDURE

The primary, methodological problem was to develop a technique which would determine how accurately Naval Aviation cadets are oriented to certain geographical points, since no standardized tests were available. A modification of the pointing procedure previously devised by High (*14*) was used because it was simple, involved a minimum of extraneous factors, was readily adaptable to group testing, and the results could be treated quantitatively.

The test, which was finally developed, consisted of two parts containing twenty items each. Each part in turn consisted of a list of ten well-known cities in the United States and ten well-known foreign cities (Tables I and II). In Part I, the subjects were asked to imagine that they were in Chicago, Illinois. Using a compass rose as a reference, they were instructed to indicate the direction they would travel from Chicago to each of the twenty cities by the shortest possible route. The directions were: "Write in Column C below the direction you would go *from Chicago, Illinois,* to reach each city by the shortest possible route. Refer to the compass rose for the directions." They read the heading from the compass rose in degrees and recorded it in the appropriate column on the test sheet. The directions for Part II were identical with those for Part I, except that the point of origin was Pensacola, Florida. These cities were selected on the basis of High's study and a preliminary study at Pensacola, eight of them being common to both points of origin. For each part, the cities were selected so that they were well distributed around the compass rose. All of the subjects were asked to face north during the test because the direction one faces during such a test has been found to be of importance in earlier studies (*14*, *16*).

The test was designed to obtain data for the analysis of errors on individual cities and to give a total orientation score. The scoring procedure used to obtain the total score on this test was somewhat unconventional. First, the latitude and longitude of each of the cities were obtained from Goode's atlas (*6*) for computational purposes. Second, initial great circle headings for each city were calculated using a method devised by Ageton (*1*). This heading is the direction a person would take to go from one point on the earth's surface to another by the 'shortest possible route'. Such headings are used by ships and aircraft when great circle sailing is followed to actually use the shortest possible distance between two points on a sphere. Third, the mean and standard deviation of the cadets' estimates for each of the forty cities were calculated. Fourth, keys were made out for each item on the test in terms of one-half sigma deviations from the initial great circle heading which was the correct response. Deviations from this heading within plus or minus one-half sigma were scored zero, and each succeeding one-half sigma step was scored as a unit, the direction of the deviation being disregarded. The sum of these deviations for all of the forty items was then used as the individual's score on the test. This score was an "error score" inasmuch as it was the sum of the individual's deviations from the initial great circle heading for each city.

Mercator maps, which plot the earth as a flat rather than a spherical surface, are commonly used in schools and elsewhere for convenience. Such projections of the earth involve gross errors, especially near the poles. However, they are widely employed so headings using Mercator plots of the cities were also determined for comparative purposes (Tables I and II). For some cities the two headings are only slightly different, but for others the errors are considerable as in the case of Moscow and Rome (Table I).

In order to compare topographical orientation with other psychological traits, additional test scores were obtained from the students' training jackets. Two additional tests of spatial ability were also given to the cadets in addition to the tests of topographical orientation which were administered during the Naval Aviation cadet's indoctrination week. The tests were:

1) *Essentials of Mathematics for Flight Students.*—This test was prepared by the Central Examining Board of the Naval Air Training Command. It is a ninety-minute test containing one

hundred and five items which varied in content from simple arithmetic computation to the reading of graphs and arithmetic reasoning given to flight students during their indoctrination week. The test is essentially a power test, since the time limit is very liberal. The standard scores obtained routinely on this test and recorded in the training records were used. No data on the reliability of the test are available.

2) *Aviation Classification Test.*—This is a group test of general intelligence used in the original selection of the Naval Aviation cadets, and was designed originally to predict success in ground school training. It consists of one hundred and twelve items measuring vocabulary, comparison, practical judgment, and arithmetic. The testing time is forty-five minutes. The reliability of the test has been found to be approximately .80.

3) *Mechanical Comprehension Test.*—This group test is also used routinely in the original selection of Naval Aviation cadets. It is essentially a test of elementary physics and mechanical relationships consisting of seventy-six items with a time limit of forty-five minutes. It was designed originally as a predictor of success in flight training. The test-retest reliability has been found to be approximately .75.

4) *Test of Correctness and Effectiveness of Expression*—This is Test I, College Form A, of the United States Armed Forces Institute Tests. There are two parts: The first emphasizes spelling and the second contains material on grammar, style, best words and the like. The time for this test is two hours and the score is the number of items correct. No data are available on the reliability of this test.

5) *American Council on Education, Psychological Examination for College Freshman,* (1947).—Standard scores were obtained from the students' records for Q (Non-language), L (Language) and total score. The standard scores used were based on a group of midshipmen having completed two years of college work.

6) *Spatial Orientation.*—This test which is No. 5 in the Guilford-Zimmerman Aptitude Survey (*13*), was designed primarily to measure the ability to appreciate spatial relations of things with reference to the human body. The test contains sixty-five multiple-choice items in which the subject is asked to make certain judgments of the position of the prow of a boat with respect to an

horizon and other objects in the visual field. The working time on the test is 10 minutes, and the reliability given in the manual is 0.88.

7. *Spatial Visualization.*—This is No. 6 of the Guilford-Zimmerman Aptitude Survey, and is designed to measure processes of imagining movements, transformations, or other changes in visual objects. The subject is asked to make judgments of the position of a clock which may be rotated about its three major axes. There are sixty-two items with a working time of thirty minutes.

The following personal data were obtained during the regular testing period: 1) age, 2) home city and state, 3) flight experience, 4) educational level, 5) travel in the United States (using a five-point check list), 6) travel outside of the continental limits of the United States (using a five-point check list), 7) geography courses taken within the last three years. In addition to this the cadets were asked to give an estimate of their own orientation using a five-point scale as follows: 8) orientation to the major points of the compass, 9) orientation during travel, 10) attention to the direction of travel in a strange area, 11) use of 'mental maps', 12) visualization of the earth (Table IV).

All of the tests with the exception of the ACT and MCT were given during the first few days at Pre-Flight School. These tests were part of an extensive battery given for experimental use or for use in the Training Command. The data obtained from these tests were coded and punched into IBM cards for computational purposes. The number of individuals used for computational purposes varies somewhat throughout the report since data were not available for all cadets on all tests, however, the specific N's are indicated in each case.

The subjects were two hundred and forty-two Naval Aviation cadets, the majority of whom were entering the Navy from civilian life; the remainder were entering from the enlisted ranks of the Navy and Marine Corps. They had passed the Naval Aviation standards on the Aviation Classification Test (above), a Mechanical Ability Test (above), and a Biographical Inventory. Eighty-five per cent of the group had completed at least two years of college. The age range was from nineteen to twenty-seven years with a mean of 21.57 and a sigma of 1.47 years.

RESULTS

(A) Orientation to Specific Geographical Points

The mean and median estimated headings for the several cities are listed in Tables I and II. The distributions of the cadets' estimates of the twenty cities for each point of origin were typically skewed to a moderate degree. In every case the range of estimated

CITY	INITIAL GREAT CIRCLE HEADING	MERCATOR HEADING	N	MEAN ESTIMATED HEADING	MEDIAN ESTIMATED HEADING	STANDARD DEVIATION	INDEX OF SKEWNESS	MEAN DIFFERENCE (Col 2-5)	CRITICAL RATIO	MEAN DIFFERENCE (Col 3-5)	CRITICAL RATIO
New York City	88	98	239	84.74	89.10	14.35	.91	3.26	3.56	13.26	14.26
Pensacola	178	179	233	176.40	168.60	12.40	1.89	1.60	1.98	2.60	3.21
Miami	157	163	240	166.08	155.40	13.38	2.39	9.08	10.53	3.08	3.57
San Diego	258	244	240	253.85	250.32	11.85	.89	4.15	5.42	9.85	12.87
New Orleans	190	189	238	187.20	184.95	12.79	.53	2.80	3.39	1.80	2.16
Seattle	295	289	238	286.52	284.95	11.95	.39	8.48	10.95	2.48	3.23
Reno	256	263	239	263.99	261.70	14.02	.49	7.99	8.81	0.99	1.09
Dallas	222	212	242	215.08	218.79	16.80	.66	6.92	6.38	3.08	2.79
Atlanta	161	166	241	166.30	152.50	20.90	1.98	5.30	3.88	0.30	16
Portland, Me.	76	80	237	73.17	66.80	14.34	1.33	2.83	3.08	6.83	7.39
London	47	77	238	86.92	81.00	18.60	.95	39.92	33.17	9.92	8.16
Tokyo	323	266	235	265.91	277.60	23.83	1.47	57.09	36.79	0.09	05
Panama	165	169	240	189.88	186.82	18.65	.49	24.88	20.60	20.12	16.74
Capetown	108	142	233	132.95	132.70	22.60	.03	24.95	16.79	9.05	6.08
Moscow	29	75	242	16.19	12.30	49.30	.24	12.81	4.05	58.81	18.45
Mexico City	207	201	240	197.29	204.35	16.90	1.26	9.71	8.83	3.71	3.41
Anchorage	320	310	236	325.12	319.02	15.20	1.20	5.12	5.22	15.12	15.15
Rome	52	90	237	100.76	105.05	18.99	.68	48.76	39.55	10.76	8.78
Rio	138	150	238	173.56	170.24	15.94	.62	35.56	34.41	23.56	22.79
Oslo	36	64	232	74.75	64.70	21.10	1.43	38.75	28.02	10.75	7.75

TOPOGRAPHICAL ORIENTATION WITH POINT OF ORIGIN CHICAGO – PART I

Table 1

CITY	INITIAL GREAT CIRCLE HEADING	MERCATOR HEADING	N	MEAN ESTIMATED HEADING	MEDIAN ESTIMATED HEADING	STANDARD DEVIATION	INDEX OF SKEWNESS	MEAN DIFFERENCE (Col 2-5)	CRITICAL RATIO	MEAN DIFFERENCE (Col 3-5)	CRITICAL RATIO
New York	43	41	239	26.32	29.74	11.97	.86	16.68	21.54	14.68	19.07
Chicago	358	359	238	353.31	351.45	12.84	.43	4.69	5.71	5.69	6.78
Miami	125	132	241	142.50	134.79	18.50	1.26	17.50	14.68	10.50	8.84
New Orleans	260	259	237	274.07	269.25	12.25	1.17	14.07	17.70	15.07	18.96
Atlanta	35	33	235	36.75	39.65	30.15	.28	1.75	.89	3.75	1.84
Denver	308	308	239	301.61	300.23	14.64	.27	6.39	6.75	6.39	6.82
Corpus Christi	256	251	237	260.65	264.15	12.00	.88	4.65	5.96	9.65	12.32
Seattle	313	310	236	305.55	310.94	12.00	1.35	7.45	9.55	4.45	5.70
Cleveland	20	20	240	9.44	6.30	15.72	.60	10.56	10.41	10.56	10.39
Lincoln, Neb.	318	330	238	332.00	323.07	17.55	1.53	14.00	12.31	2.00	1.69
London	43	66	237	72.25	62.55	16.13	1.80	29.25	27.93	6.25	6.01
Tokyo	323	273	237	296.31	288.80	23.47	.96	26.69	17.56	23.31	15.25
Honolulu, T.H.	280	260	234	275.31	274.30	15.78	.19	4.69	4.55	15.31	14.84
Panama	160	162	238	189.92	194.60	24.10	.58	29.92	19.16	27.92	17.86
Mexico City	227	223	241	234.76	234.25	25.25	.06	7.76	4.77	11.76	7.29
Coro	131	90	231	90.13	97.80	20.19	1.14	40.87	30.70	.13	.15
Toronto	23	20	241	11.40	13.51	14.50	.45	11.60	12.42	8.60	9.15
Manila	325	263	235	259.86	267.92	20.43	1.18	65.14	48.97	3.14	2.30
Reykjavik	30	36	237	31.51	32.70	15.08	.24	1.51	1.54	4.49	4.62
Buenos Aires	155	158	237	180.77	180.45	19.10	.05	25.77	20.79	22.77	48.33

TOPOGRAPHICAL ORIENTATION WITH POINT OF ORIGIN PENSACOLA – PART II

Table 2

headings was extensive as has been found with other groups (*14, 31*). Errors made by individuals approached 180° (i.e., the maximum possible error) in several cases, e.g., Chicago to Cape Town. More accurate means and particularly standard deviations were obtained by eliminating the atypical cases which it was believed were the result of errors in recording or reading. This was done by making an initial distribution for each city, determining the standard deviation and then eliminating all of those judgments which deviated more than three standard deviations above or below the mean. Therefore, the data listed in Tables I and II include only those persons who deviated three standard deviations or less from the mean. The number of eliminations in each case can be determined by subtracting the N in the table from two hundred forty-two, which was the total number of cadets tested. The number of eliminations varied from zero (Chicago to Dallas and Moscow) to eleven (Pensacola to Cairo). The major effect of this procedure was to attenuate the standard deviation while the mean was only slightly affected.

Comparison of cadets' judgments with the calculated initial great circle heading.—Deviations of the group from the initial great circle headings will give a measure of the orientation of the group to these geographical points considering the earth as a sphere. The data are uniform in showing gross individual and group deviations from the true directions (Tables I and II). Significant differences (C.R. > 3) between the great circle headings and the group mean headings occurred in thirty-seven out of the forty judgments. The deviations varied from 1.60° (Pensacola) to 57.09° (Tokyo) with the origin at Chicago and from 1.5° (Reykjavik) to 65.1° (Manila) with the origin at Pensacola. The errors were substantially less for points within the continental limits of the United States. For these points, the mean of the judgments deviated from the true direction for different cities by amounts which varied from 1.60° (Chicago–Pensacola) to 17.50° (Pensacola–Miami) with a mean of 7.51°, whereas for cities located abroad, the deviations varied between 1.51° (Pensacola–Reykjavik) to 65.14° (Pensacola–Manila) with a mean of 27.04°. The difference between the mean of the errors within the United States and abroad were significant beyond the .01 level of confidence. The errors to points within the United States were less with the origin at Chicago than when the origin was Pensacola ($P = .05$). On the other hand, there was little dif-

ference in the errors between the two points of origin for cities located abroad. These data show that there are constant tendencies within the cadet group to misjudge the true directions of cities within the United States and abroad. However, these group errors are much less than the errors made by individuals. Some illustrations will be useful in showing the size of these errors: New York is almost due east of Chicago, but several people indicated that it is northeast, while several others indicated that it is southeast of Chicago. From Pensacola to Panama, the initial great circle heading is approximately southeast, while the estimated direction by nearly half of the cadets was southwest. Very large errors were also found from Chicago to Tokyo, Moscow, Rome, and Oslo, and from Pensacola to Cairo, Manila, and Tokyo.

Comparison of cadets' judgments with headings determined from Mercator projection maps.—A major problem which has confronted map-makers for many years is how to make a satisfactory projection of the spherical surface of the earth on a plane surface. All of the methods which have been devised to do this result in certain distortions of shape, scale, angle, and area. The Mercator projection has been widely used in atlases in the past notwithstanding the fact that it produces gross distortions of size as the poles are approached. Deviations of the cadets' judgments from the direction of the cities as determined from Mercator projections were calculated as in the preceding section (Tables I and II). The mean of the deviations of the cadets' judgments from the Mercator heading were substantial. However, these deviations were less than the deviations from the initial great circle headings. Differences between the mean of the cadets' judgments and the Mercator heading between the cities exhibited critical ratios greater than three in thirty-two out of the forty judgments. The mean of the deviations for points within the United States was less for the Mercator projections than for the great circle headings, but this difference was not statistically significant. The deviations from the Mercator headings for cities outside of the continental limits of the United States were significantly less than those for the great circle headings ($P = .02$). The deviations varied from 0.09° (Tokyo) to 23.56° (Rio de Janeiro) with the origin at Chicago and from 0.13° (Cairo) to 27.92° (Panama) with the origin at Pensacola. For cities within the continental limits of the United States, the smallest deviation for the group was 0.30° (Chicago–Atlanta),

while the greatest was 15.07° (Pensacola–New Orleans). The mean deviations from Mercator headings within the United States was significantly less than for cities abroad (P = .03). These data show clearly that gross deviations from the Mercator headings occur both for the group and for individuals within the group. It is also evident that the cadets are more closely oriented to the earth as a flat surface than to the earth as a sphere.

Discussion.—These results demonstrate that many Naval Aviation cadets are very poorly oriented to certain geographical points. They also show that there are large group errors as well as individual errors. The results are in general agreement with other studies which have found great variability from person to person, and substantial variation from place to place with regard to group errors (*14*, *31*, *32*). However, these results are directly opposed to Woodring's (*32*) reports of bimodal distributions. His subjects appeared to fall into two groups, one of which was well oriented and the other making approximately a chance score. No clearly defined bimodal distributions were found in the forty cities studied here, although the distributions were skewed. The explanation of this difference is to be found in the fact that the cadet population was larger and much more homogenous than the group studied by Woodring. These cadets were all men of a limited age range, having high aptitude test scores and, in general, extensive academic background. On the other hand, Woodring's group was highly heterogenous. "Of the one hundred fifty-six subjects, one hundred and seven were white and forty-nine were Negroes; one hundred and seventeen were males and thirty-nine were females. The age range was from seventeen–sixty. Intelligence quotient—by the Revised Stanford Binet Scale—ranged from 44–122. The educational range was from no formal education to college graduation." (*32*) Since Woodring found several of these factors to be related to orientation, it is conceivable that selective factors may have contributed to the bimodal distribution in his population.

The results of this study are in agreement with High's (*14*) finding that foreign cities show greater errors than cities within the United States and Woodring's finding that distant cities show greater errors. It is interesting to note, however, that High (*14*) found great variability for points within the city where his subjects took the test. This may have been a function of the selection of his points and the fact that some of his subjects were not thor-

oughly familiar with the city. However, High's data suggest that a person may be poorly oriented in a limited area with which he is fairly familiar but well oriented in a larger area. Additional comparisons with High's data are possible since ten of his cities were identical with the ones used in this study for the point of origin at Chicago. The cadets were somewhat more accurate and less variable than the college group. However, the differences between the groups do not appear to be great in the light of the differences between the groups and the test situations. There was no significant difference (C.R. < 3) for four of the cities (New York, Panama, Cape Town, Moscow) while significant differences were found for six cities (Miami, New Orleans, London, Tokyo, San Diego, and Seattle).

The data show that the orientation of the cadets at the beginning of training is much closer to Mercator projections than to great circle or global projections. This is true in spite of the fact that the deviations are large in both cases. It would appear that the wide use of Mercator and other similar flat projection maps in school and elsewhere has had a marked effect on the cadets' orientation to points on the earth producing marked disorientation, particularly to foreign cities. Greater emphasis on the relation of geographical points to each other on a globe in schools and elsewhere would appear highly profitable to reduce or eliminate such gross errors.

No attempt was made in this study to analyze the mechanism used to make the judgments of the direction of the cities. However, certain inferences may be drawn from the data at hand in terms of the type of errors made. Firstly, it is evident that the cadets as a group have acquired certain erroneous associations in terms of distances and/or direction. For example, the cadets as a group believe that Panama is well west of Chicago and Pensacola when it is actually east of these cities. Secondly, the cadets' orientation is much closer to Mercator projections than to a global map. Indeed, they show little appreciation of great circle courses between points and apparently have little appreciation of the location of cities on the world as a sphere. An illustration of the very large errors which result from this type of disorientation is to be found in the group deviation of 65.14° from the great circle heading from Pensacola to Manila. Although Manila is slightly south of Pensacola, the great circle heading would be roughly north-northwest because the

two cities are well north of the equator. However, the mean of the cadets' judgments for the shortest route was slightly south of due west (Table II). The deviation of this direction from the Mercator heading was only 3.14° which was not significant statistically. In other words, the cadets as a group are fairly well oriented to Manila from the point of view of Mercator projections, but they are thoroughly disoriented from a global point of view. Similar gross deviations from great circle headings with small deviations from Mercator headings are found from Pensacola to Cairo and London and from Chicago to London, Rome and Tokyo (Tables I and II). It is interesting to note that the deviations from the Mercator heading are greater from Chicago to Moscow. The cadets tend to indicate a heading close to the North. This may be explained as an artifact resulting from the instructions which point out that the Great Siberian Plain can be reached most directly by passing near the North Pole. These results give an explanation for the fact that foreign cities show greater errors than cities within the United States. Mercator projections are sufficiently accurate within the United States to produce only small errors, whereas Mercator projections are inadequate for cities abroad. Foreign cities like Toronto and Mexico City which are fairly well represented by the Mercator projection show errors of similar amounts for both Mercator projection and great circle heading.

(B) The Relationship of Topographical Orientation to Other Psychological Factors

The intercorrelations between the scores on the topographical orientation test and other psychological factors are summarized in Table III. Most of these correlations are not significantly different from zero, and all of them are extremely low, indicating that performance on the topographical orientation test is not closely related to performance on the other tests given. It should be noted that the topographical orientation score is an error score, since it was the sum of the individual's deviations from the initial great circle heading for each city. (The odd-even reliability of this test was: Part I—.59; Part II—.65; and Total Score .75, when corrected by the Spearman-Brown formula, the reliability of the total score becomes .86). The test of topographical orientation showed the highest relationship to the Q score on the American Council on Education, Psychological Examination and the test of Mathe-

TABLE III.—Correlation between the Test of Topographical Orientation and Other Variables. (Note that the topographical orientation score is an error score.)

	Topographical Orientation			N
	Part I	Part II	Total Score	
1) Spatial Orientation (Guilford-Zimmerman, Part V)	.06	−.01	−.04	233
2) Spatial Visualization (Guilford-Zimmerman, Part VI)	−.13*	−.07	−.11	233
3) Aviation Classification Test	−.06	−.10	−.09	224
4) Mechanical Comprehension Test	−.14*	−.02	−.08	224
5) American Council on Education—Q	−.19†	−.19†	−.21†	213
6) American Council on Education—L	−.04	−.08	−.07	213
7) American Council on Education—Total	−.15*	−.17*	−.18†	233
8) Correctness of Expression	−.15*	−.14*	−.16*	233
9) Mathematics	−.18†	−.21†	−.21†	232
10) Personal Estimate of Orientation	−.19†	−.09	−.15*	233
11) Age	.01	.06	.04	231
12) Educational Level	.04	.08	.07	233

* r −.13 significant at .05 level.
† r −.18 significant at .01 level.

matics. There is a less significant correlation between topographical orientation test scores and the cadets' own evaluation of their orientation and the test of correctness and effectiveness of expression.

No significant relationship was found between the test of topographical orientation and the test of spatial visualization, which is contrary to the results found by High (*14*), who reported an *r* of .51 for his group. The reason for the discrepancy between High's results and the results of this study may be found in the fact that (a) High's group contained only seventy-five individuals, and (b) High's group was much more heterogeneous, containing both men and women college students. It would appear, therefore, that in a cadet population, spatial visualization and spatial orientation are not related to topographical orientation, as measured by this test.

The data for Naval Aviation cadets confirm Woodring's findings that there is no relation between topographical orientation and age or intelligence; this is true notwithstanding the fact that Wood-

ring's population was highly heterogeneous (*31*, *32*). As has been pointed out above, Woodring also found for his subjects (based on six geographical points) that the distribution of average scores was bimodal and the individuals could be classified into three fairly distinct groups. No evidence of bimodality was found for the individual cities as noted in the preceding section. Similarly, the total score for the topographical orientation test gives no evidence of bimodality; however, the scores do fall on a skewed distribution. For the group of two hundred and thirty-three persons used in this part of the study, the mean was 78.04 with a standard deviation of 33.40 and a range from 40 to 390. However, only four of these scores fell between 187 and 390, while only six scores fell between 138 and 179. If these ten scores are eliminated, the mean is slightly reduced to 73.19, while the standard deviation is greatly reduced to 16.80. The remaining scores show only slight skewness (index skewness equals .30) toward poor orientation as compared with a moderate index of skewness (.50) for the total population (*12*). The distribution of scores on the topographical orientation test can be described as being moderately skewed toward poor orientation, but there is no indication of bimodality in these data.

A composite personal orientation score was obtained for each of the subjects (items 5–12, Table IV) from the personal data listed in the procedure. Each of these eight items was assigned a weighted score depending upon which of the choices listed under that question was checked. The scores for the individual items were then summed and this total was assigned as a measure of the personal orientation. To determine whether there was a relationship between the answers to each of the questions and scores on the topographical orientation test, the following procedure was followed for each question: 1) the responses were sorted into like groups, i.e., all subjects who marked question 5, 'chiefly near home town', were placed in one group; 'chiefly in home state' in a second group, etc. 2) The scores on the topographical orientation tests for each of these groups were obtained and an analysis of variance was performed to determine where the differences between group performances were significant. 3) If a significant F was obtained, thus indicating that several groups differed significantly among themselves, a *t* test was run to determine those differences between means which were significant.

Questions 4, 5, 9, 11 and 12 provided F's significant at the .001

TABLE IV. SUPPLEMENTARY DATA

Name ..... 1. Age (nearest year).....

2) Home city: ....., State: .....

3) Flight status: ..... aviation training, ... stage; .....pilot, ... nonpilot; approximate hours of flight time .....

4) Encircle the highest educational level finished: High school—1, 2, 3, 4 years; College—1, 2, 3, 4 years; Post graduate.

5) Travel in the United States: ..... chiefly near home town, ... chiefly in home state, ... chiefly in home and nearby states, ... one or two transcontinental trips (or trips over 1,500 miles), ... three or more transcontinental trips (or trips over 1,500 miles).

6) Travel outside the continental limits of the United States: ..... never, ... one or two visits to Canada or Mexico, ... several trips to Canada or Mexico, ... one trip overseas, ... two or more trips overseas.

7) How many courses in geography or related subjects have you had in the last three years? .....

8) Do you keep the direction of the major compass points (N, S, E, W) in mind when riding as a passenger in a car?

Always	Very often but not always	Sometimes if I have a reason	Occasionally when not familiar with the terrain	Almost never

9) Do you usually orient yourself in travel by means of some reference point, such as a city or some other place whose location is familiar to you; i.e., can you point to it if you wish?

Never	Occasionally if I have a special reason	Sometimes if possibility of becoming lost	Quite frequently	Always out of habit

10) If you are traveling in a strange area (city or country), do you try to remember the number and direction of the turns you have made so that you will be able to find your way back the same way if necessary?

Almost always	Frequently	Sometimes if I have a special reason	Occasionally if I *know* I will have to return the same way	Never

TABLE IV.—*Continued*

11) Do you keep a "mental map" of the area you are in, i.e., if you were required to do so could you reproduce a rough approximation of the area in which you find yourself as by drawing a crude map?

Never; I am not concerned about these things	Only if the area is well known to me	I do this as much as most people	Very often	Invariably through habit

12) Can you visualize a globe of the world with the continents and major bodies of land in their approximate positions with respect to each other?

Confident I have an accurate picture of globe	Believe I know positions of continents very well	My picture of globe is about average	Unsure of shape and position of some continents	I have a very poor picture of globe in mind

level and subsequently the differences between the means of the subgroups for each question were examined to determine those differences which were significant. For question 4, the group with a high school education was significantly better oriented (P = .01) than the group with one year of college. For question 5, the group which had answered, 'chiefly in home town', had poorer orientation (P = .01) than the groups which had answered the remaining four responses. For question 9, no significant differences were obtained between the means of the groups; this tends to indicate that although the samples were not drawn from a common population the difference is in variances and not in means. For question 11, the group which had answered, 'invariably through habit', showed better orientation (P = .05) than the groups which had answered, 'very often' and ' I do this as much as most people'. For question 12, the group which answered, 'My picture of the globe is about average', showed poorer orientation (P = .01) than the groups which answered, 'confident I have an accurate picture of globe' and 'believe I know position of continents very well'. Also, the cadets who answered, 'unsure of shape and position of some continents', showed poorer orientation (P = .05) than the cadets who answered, 'confident I have an adequate picture of globe'.

The choices for questions 8 through 12 were examined to deter-

mine how closely the answers approximated the distribution of individuals normally expected to fall within the various classes of ratings (*12*); assuming that the trait that is rated has a normal distribution in a fairly large unselected group. None of the questions met the criterion. It must be pointed out that this lack of continuity of the items of these questions no doubt had an important effect upon the variable results which were obtained. It is recognized that this weakness must be remedied in any future use of a personal orientation questionnaire. For this reason, a minimum amount of emphasis is placed upon the results previously mentioned. However, it is also noted that the results obtained point favorably to the feasibility and utility of such a revised personal biographical instrument to augment any future studies on topographical orientation.

Discussion—The data show that the factors measured by the topographical orientation test are unique. Very low correlations were found between these scores and those on two other orientation tests, two intelligence tests and a mechanical aptitude test. A similar result was found for a variety of personal data, though some individual items of the personal questionnaire did contain items which were significantly related to topographical orientation. The fact that those who indicate travel, 'chiefly near home town', show poorer topographical orientation than those with more extensive travel experience might have been anticipated. However, it is interesting to note that the other categories show no significant differences. The significantly better orientation of individuals with no more than high-school education as compared with those with college training tends to support Woodring's (*32*) finding that persons with lesser amounts of schooling do better on such tests. The non-college cadets come from a group which has had a relatively high amount of travel experience in the Navy which is undoubtedly an important factor. It is apparent that the current selection devices in the Navy do not include a variable related to topographical orientation as measured by this test. The relation of this composite score to success in Naval Aviation Training is not known. However, a study currently underway will obtain data on this point.

It will be noted that all of the deviations measured by this test were measured from initial great circle headings which would give the correct initial heading to reach the city by the shortest possible

route. Similar calculations were made using the mean heading indicated by the group. This attenuated the error scores, but it did not change the intercorrelations significantly. Scatter plots of the data also indicated that the elimination of the cadets having the most widely deviant scores would not cause any significant change in the correlations.

SUMMARY

A test of topographical orientation and seven other tests were given to two hundred and forty-two Naval Aviation cadets to study the types of errors in orientation they make, to identify the factors which contribute to this type of disorientation, and to determine what relation there is between topographical orientation and certain other psychological factors. In the test of topographical orientation, large errors were found to be the rule for orientation to individual geographical points. Deviations of the group from proper orientation were as great as 65.1°, while some individual deviations from the proper orientation approximated 180°, the maximum possible error. Errors for cities located abroad were greater than for cities within the United States. Deviations from Mercator headings were large, but these errors were much less than for deviations from great circle headings, particularly for cities outside of the continental limits of the United States. Correlations between total score on the test of topographical orientation and other psychological tests were .21 or below; only three of them being significantly different from zero. Naval Aviation cadets are poorly oriented to geographical points, particularly as they actually occur on the globe. The notion of great circle relationships between distant cities does not appear to enter significantly into their judgments of the geographical points. Topographical orientation appears to be a unique characteristic which is independent of the factors measured by the current selection battery for cadets as well as certain other factors such as mathematics, spatial orientation, and spatial visualization.

REFERENCES

1) A. A. Ageton. *Dead reckoning, Altitude and Azimuth Table.* Washington, D. C.: U. S. Navy Hydrographic Office, No. 211, U. S. Government Printing Office, 1940.

2) A. Binet. "Reverse illusions of orientation." *Psychol. Rev.*, 1894, 1, 337–350.

3) H. R. De Silva. "A case of a boy possessing an automatic directional orientation." *Science*, 1931, 73, 393–394.

4) F. N. Freeman. *Psychology of the Common Branches*. Boston: Houghton-Mifflin Co., 1916. (Chapter VIII)

5) J. J. Gibson. *The Perception of the Visual World*. Boston: Houghton-Mifflin Co., 1950.

6) J. P. Goode. *School Atlas*. New York: Rand McNally and Co., 1943.

7) T. Gordon. *The airline pilot: a survey of the critical requirements of his job and of pilot evaluation and selection procedures*. Washington, D. C.: C.A.A. Div. of Research, Report No. 73, 1947.

8) A. Graybiel. "Spatial disorientation in flight." U. S. Naval School of Aviation Medicine, Research Report No. NM 001 059.01.24. (also in The Military Surgeon, 1951, 108, 287–294).

9) D. R. Griffin. "The sensory basis of bird navigation." *Quart. Rev. Biol.*, 1944, 19, 15–31.

10) —— "Topographical orientation." (Ch. 17 in C. G. Boring, et al, *Foundations of Psychology.*) New York: Wiley, 1948.

11) D. R. Griffin and R. J. Hock. "Experiments on bird navigation." *Science*, 1948, 107, 347–349.

12) J. P. Guilford. *Psychometric Methods*. New York: McGraw-Hill, 1936.

13) J. P. Guilford and W. S. Zimmerman. *The Guilford-Zimmerman Aptitude Survey*. Beverly Hills, California: Sheridan Supply Company, 1947.

14) W. S. High. *An Investigation of Geographical Orientation and Its Relation to Spatial Visualization*. Unpublished M.A. thesis, San Jose State College, San Jose, Calif., 1951.

15) G. F. Howe. "The teaching of directions in space." *J. of Geog.*, 1931, 30, 298–304.

16) —— "The teaching of directions in space." *J. of Geog.*, 1932, 31, 207–210.

17) J. I. Lacey and S. W. Niehaus. "Orientation tests" in J. P. Guilford *Printed Classification Tests, Report No. 5*, Army Air Forces Aviation Psychology Program Research Reports. Washington, D. C.: U. S. Government Printing Office, 1947.

18) F. E. Lord. "A study of spatial orientation of children." *J. ed. Res.*, 1941, 34, No. 7, 481–505.

19) E. Meyer. "Comprehension of spatial relations in pre-school children." *J. genet. Psychol.*, 1940, 57, 119–151.

20) W. B. Michael, W. S. Zimmerman, and J. P. Guilford. "An investigation of two hypotheses regarding the nature of the spatial relations and visualization factors." *Ed. and Psychol. Measurement*, 1950, 10, 187–213.

21) K. C. Montgomery and E. G. Heinemann. "Concerning the ability of homing-pigeons to discriminate patterns of polarized light." *Science*, 1952, 116, 454–456.

22) H. E. Page. "Vertigo incidence among naval aviators." U. S. Naval School of Aviation Medicine, Research Report No. NM 001 063.01.16, 15 March 1950.

23) D. Ridgley. "The teaching of directions in space and on maps." *J. of Geog.*, 1922, 21, 66–72.

24) W. Rowan. "Homing, migration and instinct." *Science*, 1945, 102, 110–111.

25) T. A. Ryan and M. S. Ryan. "Geographical orientation." *Amer. J. Psychol.*, 1940, 53, 204–215.

26) L. F. Shaffer, et al. "Surveys of experiences of returned personnel," in F. Wickert, Ed., *Psychological Research Problems of Redistribution*, Army Air Forces Aviation Psychology Program, Research Report No. 14. Washington, D. C.: U. S. Government Printing Office, 1947.

27) C. C. Trowbridge. "Fundamental methods of orientation and imaginary maps." *Science*, 1913, 38, 97.

28) W. E. Vinacke. "Illusions experienced by aircraft pilots while flying." *J. Aviat. Med.*, 1947, 18, 308–325.

29) H. C. Warren. "Magnetic sense of direction." *Psychol. Bull.*, 1908, 5, 376–377.

30) H. A. Witkin. "Studies in geographic orientation." *Yearbook Amer. Phil. Soc.*, 1946, 152–155.

31) P. D. Woodring. "A technique for the investigation of direction orientation in human beings." Michigan Acad. Arts and Sciences, 1938, 24, Pt. IV, 147–152.

32) ——. "An Investigation of Direction Orientation in Human Adults." Ab. Doctoral Dissertations, No. 29. Columbus: Ohio State Univ. Press, 1939.

STUDENT DIFFERENCES IN COMPETENCY IN EDUCATIONAL PSYCHOLOGY PRIOR TO FORMAL TRAINING[1]

WILLIAM W. LYNCH, JR.

Indiana University

The research to be reported here is an attempt to analyze and study two types of competence that are commonly stressed in the training of prospective teachers in educational psychology: (1) familiarity with the principles, facts, and value judgments most typically presented in textbooks in educational psychology, and (2) ability to interpret information about an individual child's behavior problems, as presented in written case-study form, and make valid recommendations for the solution of these problems. The need for such analysis can be supported by a review of the literature of the past several decades dealing with the content and teaching of educational psychology. In recent years, particularly, the discussions of ways and means of improving instruction in educational psychology have stressed the opinion that the psychological competencies needed by teachers are not fulfilled by familiarity with principles, facts, and authoritative judgments, and that this familiarity and its applicative concomitants are inextricably involved with the prospective teacher's total personality make-up. Just what the nature of these relationships may be is far from clear. It was the intent of the present study to probe this complex of relationships by studying the two abilities mentioned above as measured in a population of teachers college students planning to become elementary teachers who had received no formal instruction in psychology. By selecting such an untrained group it was possible to discern the extent of differences in these two abilities prior to training and to study the relationship of these differences to several psychological characteristics and background experiences of a primarily non-academic nature. The related factors selected were: tested intelligence; personal and social adjustment; occupational interests; attitudes toward child care and discipline; and amount of past experience in caring for, teaching, or directing children.

[1] Based in part upon the writer's dissertation submitted for the Ph.D. at Yale University, 1950.

HYPOTHESES TO BE TESTED

On the basis of a review of experimental studies and published authoritative opinion, and a logical analysis of the factors involved in the study, the following seven hypotheses were formulated for testing:

I) Students who have had no instruction in educational psychology differ significantly in their familiarity with the facts and principles of this field that are contained in textbooks. Some of the facts and principles are familiar to a majority of these students.

II) Students who are most familiar with the content of educational psychology are best able to interpret case-study data and to select valid recommendations for remedial measures.

III) Among persons with no formal training in psychology there is a positive relationship between tested intelligence and both (1) familiarity with facts and principles and (2) ability to interpret case study data correctly.

IV) Prospective teachers with no formal psychological training who have had extensive experience with young people are better informed in the content of educational psychology and are more competent to interpret case material.

V) Prospective teachers prior to psychological training who have measured interests in social activities and in the care and direction of others are more proficient in both knowledge and applicative ability than those who show little or no interest in such activities.

VI) Good social and personal adjustment is positively correlated with the informational and applicative criteria of proficiency in educational psychology.

VII) Persons who manifest strict authoritarian attitudes toward child behavior are poorly informed in educational psychology and are relatively lacking in the ability to interpret case material.

SUBJECTS

Four hundred and forty freshmen enrolled in three Connecticut state teachers colleges that prepare elementary teachers were selected to test the hypotheses. These were divided into subgroups. The numbers in each subgroup and the hypotheses to be tested are shown in Table I. All testing was done during the first half of the fall semester of the 1949–1950 academic year.

TABLE 1. ARRANGEMENT OF GROUPS FOR TESTING

College	N	Tests	Hypotheses
A	220	Subtest on 'Personality'	I, II, III
	47	Subtest on 'Growth and Development'	
	45	Subtest on 'Individual Differences'	
	48	Subtest on 'Motivation'	
	48	Subtest on 'Learning'	
	57	Subtest on 'Problem-solving'	
	58	Subtest on 'Evaluation'	
	69	The Case of Mickey Murphy	
B	131	Abbreviated test of Information	II, III, V, VI
		The Case of Mickey Murphy	
		The Minnesota Personality Scale	
		The Kuder Preference Record	
C	93	Abbreviated Test of Information	II, III, IV, VII
		The Case of Mickey Murphy	
		Experience Rating Scale	
		Scale of Attitudes Toward Child Behavior	

CRITERION TESTS

1) *Content Tests.*—A comprehensive battery of 432 objective test items was constructed by the writer and an assistant. These items were based upon the content of fourteen textbooks in educational psychology published between December, 1945, and August, 1949. The battery was divided into seven subtests covering the following topics most commonly treated in the textbooks:

Individual differences, growth and development, personality, motivation, learning, problem-solving, and the evaluation of school achievement. The curricular validity of the battery was checked by submitting the items to four educational psychologists whose training and teaching experience presumably qualified them to judge the validity of the items. The split-half reliability of the seven subtests, corrected by the Spearman-Brown formula, ranged from .21 to .67, the median being .53. An abbreviated form of the entire battery was constructed by selecting items at random from the original subtests in proportion to the amount of space devoted to the subtest topics in the fourteen source textbooks. This

form had a reliability coefficient of .58, as corrected by the Spearman-Brown formula.

2) *Case-Study Test.*—A published case-study test, *The Case of Mickey Murphy*[2], was selected as the criterion measure of ability to diagnose a problem involving a young person in a school situation and to make valid recommendations for remedial procedures. This test was based upon actual case-study data and was validated by the judgments of experts. The odd-even reliability obtained on the group tested in this study was .64 when corrected by the Spearman-Brown formula.

DEPENDENT VARIABLE MEASURES

1) *Intelligence.*—Scores on the *California Test of Mental Maturity, Advanced '47 S-Form,* were obtained for all subjects. This test had been administered as an admissions test prior to this investigation.

2) *Experience with Children.*—A check list of activities such as baby-sitting, Sunday-school teaching, etc., was devised and accompanied by a rating scale designed to indicate the extent of experience in the care and direction of children. These instruments were based upon a preliminary questionnaire study of the experiences of a group of education and psychology students. The freshmen in this study were asked to check all of their experiences involving the care and direction of young people, and then were asked to estimate the extent of this experience when compared with that of full-time classroom teaching in terms of actual time spent in contact with children. Length-of-experience intervals ranged from 'one to three months' to 'two or more years'.

3) *Interests.—The Kuder Preference Record—Occupational* was used as a measure of occupational interests. The data from male and female subjects were treated separately.

4) *Personal and Social Adjustment.*—The *Minnesota Personality Scale* was administered to the freshmen used in testing Hypothesis VI. Male and female data were treated separately.

5) *Attitudes toward Child Behavior.*—The writer constructed a nineteen-item attitude scale based upon the written opinions of students in classes in psychology and child development on a child behavior problems such as discipline, freedom of choice

[2] Warren Baller, *The Case of Mickey Murphy,* Lincoln: The University of Nebraska, 1943.

for children in everyday matters, the use of punishment and reward, academic standards, and the characteristics of a successful public-school teacher. The scale was designed to measure on an 'authoritarian-permissive' continuum. The Likert technique for construction and scoring was used. No direct check on validity could be made, but a preliminary scale was submitted to three experienced instructors in education and psychology who were asked to indicate the responses to each item that would reveal in their opinion attitudes at each end of the postulated continuum. Items on which the judges disagreed or which were judged to be either ambiguous or not related to the postulated continuum were discarded. The split-half reliability, corrected by Spearman-Brown formula, was .58. The test-retest reliability, with an intervening interval between testings of about four months, was .65 for 34 subjects.

PROCEDURE

Hypothesis I.—The schedules of the two hundred and twenty freshmen subjects in college A precluded the possibility of administering the entire battery of content tests to each freshman. Therefore, the group was divided into five sections and three of these sections were administered three content tests each, and two sections were administered one content test and the case-study test (see Hypothesis III). All groups took one of the subtests in common so that a rough check could be made of the equivalence of the five freshman sections (Table I).

The content test scores were treated both as raw scores and as 'true per cent scores' which were based upon theoretical 'zero' scores (scores theoretically obtainable by chance). The mean 'true per cent scores' and the ranges of raw scores were compared for the seven subtests to determine whether a distinct profile of subtest scores existed among freshmen. The freshman scores were then compared with the distributions of scores obtained from seniors in the same college and the extent of overlap was determined. Finally, an item analysis for all content tests was carried out in order to determine significant differences in responses of freshmen and seniors.

Hypothesis II.—One hundred and thirty-one freshmen in college B and ninety-three freshmen in college C were tested with the 'abbreviated' content test which was derived from the total

content test battery. The same groups were also administered the case-study test. Pearson product-moment correlation coefficients between content scores and case-study test scores were computed and corrected for attenuation. The discovery of significant correlations between IQ and the two criterion tests made it desirable to compute first order partial *r*'s between the two criterion scores partialing out IQ.

Hypotheses III, V, VI, and VII.—These hypotheses were tested by computing zero-order correlations between each of the two criterion measures and the measures of intelligence, occupational interest, personality, and attitude. In each case the abbreviated content test was used.

Hypothesis IV.—Hypothesis IV was tested by determining the significance of the difference between the mean criterion scores obtained by the top and bottom twenty-five per cent of the group as determined by self ratings of length of experience.[3]

A second method of testing Hypothesis IV was devised by resorting the self-ratings together with the check list data and selecting two groups of high and low experience as determined by the combined criteria of self-ratings and the writer's judgment of the check list data.

RESULTS

1) Hypothesis I, positing the existence of differences in familiarity with the content of educational psychology among beginning prospective teachers, appears to be substantiated by the data obtained. In the first place, on both the separate topical area tests and on the total test the freshmen scored significantly above the theoretical 'zero' score obtainable by chance. In the second place, these freshman scores generally ranged from a low around the 'zero' score to a point sufficiently high to result in a mean overlap of 7.1 per cent of freshman scores above the median scores of seniors who had completed at least three courses in psychology (general, child, and educational psychology) in addition to methods courses and practice-teaching. These two general findings indicate that freshmen do differ significantly in their knowledge of the facts and principles in educational psychology. These findings are summarized in Table II.

[3] The small number of experience categories and a skewed distribution prevented a correlation treatment.

TABLE 2.—SUMMARY OF RESULTS ON CRITERION TEST OF EDUCATIONAL PSYCHOLOGY CONTENT

Group	N	Test	Theoretical Chance Score	Median	Range	Median Senior Score*
College A	47	Growth and development	24.1	33.6	17–46	38.7
	45	Individual differences	24.5	31.2	25–41	38.2
	47	Motivation	25.5	37.4	24–45	43.8
	48	Learning	24.4	35.5	29–48	39.9
	57	Problem-solving	23.0	34.6	22–43	40.5
	58	Evaluation of school Achievement	22.2	32.6	20–40	38.8
All Sections	220	Personality	22.7	33.2	20–45	41.9
Colleges B and C	224	Abbreviated form	22.0	33.0	20–45	38.9

* Scores obtained on 112 Seniors in College A, 52 Seniors in Colleges B and C.

The freshman responses to the separate items in the content battery, when compared with the senior responses, indicate general familiarity with certain facts and principles in educational psychology, at least at a recognition level. Of the four hundred and thirty-two items in the total battery, two hundred and eight were answered correctly by a significant majority of freshman subjects. Two hundred and fifty items showed no significant difference between the freshman and senior percentage answering correctly. Of the one hundred and eighty-two items on which there were significant differences between the two groups, twelve were answered correctly by a significantly larger proportion of freshmen than seniors.

It was difficult to categorize the items which most freshmen answered correctly. Generally speaking, they consisted of generalizations expressed in nontechnical language. In some cases freshman superiority over senior responses plainly indicated that general information or 'common knowledge' that is valid psychological fact had been distorted by seniors through poor comprehension of psychological data during training.[4] The mean freshman scores

[4] For example, ninety-six per cent of freshmen recognized that 'drill facilitates learning,' while only seventy-six per cent of the seniors answered

on the subtests, while showing some small differences that approached statistical significance, do not warrant a clearcut conclusion as to the relative familiarity with the seven topical areas covered by the content test battery.

A word of caution is in order in interpreting the results pertaining to Hypothesis I. No attempt was made to discern evidence of response sets which conceivably may have affected some of the item-analysis results. Secondly, it is not redundant to point out that the behavior of these subjects being measured was most probably a very low order recognition type of response and the term familiarity, used in the statement of the hypothesis, should be interpreted accordingly.

2) Hypothesis II, stating that there would be a positive relationship between familiarity with textbook content and ability to interpret case-study data, was definitely confirmed. For College B and College C freshmen the uncorrected correlations between the content test scores and the case-study test scores were .46 and .37, respectively, (both significant at the .01 level). When corrected for attentuation these coefficients were .76 and .61. Partial *r*'s were computed, holding tested intelligence constant. The partial *r*'s were .64 and .55, respectively. These results indicate that theoretically the relationship between the two variables is fairly high; somewhat higher, in fact, than those found in most investigations of the relationship between knowledge and ability to apply.[5]

3) The positive relationship between tested intelligence and the two criterion measures of psychological proficiency posited in Hypothesis III was borne out. College B and College C freshmen yielded correlations of .37 (SE_r = .08) and .43 (SE_r = .06) for the content test scores, and .39 (SE_r = .08) and .35 (SE_r = .09) for the case-study test scores. Correction for attentuation raises each of these from ten to fifteen points.

---

this correctly. Obviously there was confusion between the effect of drill upon certain learning outcomes with the undesirability of drill when applied to certain educational objectives.

[5] Maurice B. Troyer and Horrocks, John B. "Case-study tests of ability to use knowledge of human growth and development," *Educational and Psychological Measurement*, 7: No. 1, 23–26, 1947. Ralph W. Tyler, "The relation between recall and higher mental processes," Chapter II in Charles H. Judd, *Education as Cultivation of the Higher Mental Processes*. New York: Macmillan Company, 1936, pp. 6–17.

4) Two methods of selecting high- and low-experience groups from the freshmen of College C and comparing these groups' mean scores on both criterion tests yielded no differences that were statistically significant.

5) Interest test scores in 'social service' for male and female subjects yielded correlations with the criterion tests that did not differ significantly from zero.

6) Separate correlation of five subtest scores on the *Minnesota Personality Scale* with criterion test scores for male and female groups did not differ significantly from zero.

7) Hypothesis VII, which assumes a relationship between attitudes toward child behavior on an 'authoritarian-permissive' continuum, seemed to be confirmed. The correlation of raw scores on the attitude scale with the criterion tests were: content test, .24 (SE_r = .10) (significant at the .05 level but not at the .01 level); case-study test, .32 (SE_r = .10) (significant at the .01 level). These indicate a tendency for those with 'permissive' attitudes to score higher on both tests. When corrected for attenuation these correlations become .51 and .52, respectively.

These results are not surprising when an examination is made of the separate items on both tests. Recognition of the 'correct' answers on the content test frequently requires the testee to make a value judgment rather than a decision of correctness of fact or theory. The sampling of the textbook content required for this test included judgments of textbook writers concerning desirable educational practice. The same point may be made with regard to the case-study test. Here also many questions are concerned with remedial practices in a school situation. A person generally in favor of more lenient or permissive handling of educational problems would be apt to do better on either test.

SUMMARY AND CONCLUSIONS

Freshmen in three Connecticut state teachers colleges were administered tests of information based on the content of fourteen textbooks in educational psychology and a case-study test purporting to measure ability to interpret and make valid remedial recommendations concerning an educational problem of a psychological nature. Scores on these tests were examined to determine: (1) the extent of differences in ability to recognize correctly facts, principles, and authoritative opinion as contained in educational

psychology textbooks prior to any formal training in this subject; (2) the relationship between familiarity with textbook content and ability to interpret case-study data; and (3) the relationship of the two measures of psychological proficiency to measures of intelligence, experience in care and direction of young people, interest in social service, social and personal adjustment, and attitudes toward child behavior and discipline.

It was found that freshmen demonstrate some familiarity with the facts and generalizations of educational psychology prior to training and that there is a significant spread of differences in such familiarity. Content familiarity and ability to interpret case-study data are positively related to a relatively high degree.

Tested intelligence and attitudes toward child behavior are significantly related both to the knowledge and to the applicative aspects of educational psychology. Experience with children, vocational interest in social service, and personal and social adjustment as measured by the paper-and-pencil instruments used, show no significant relationship with the two criteria of proficiency.

It should be emphasized that the study was an exploratory one, using instruments whose validity and reliability were suitable for a general, theoretical examination of the attributes concerned and not for the precise testing of exact hypotheses. In spite of this general nature of the study, several implications pertaining to educational psychology in teacher education are suggested:

1) The study's confirmation of the existence of considerable heterogeneity of psychological proficiency among prospective teachers prior to formal training in educational psychology reinforces the general principle that instruction in this field, as in others, needs to be geared to the differences that exist. Furthermore, not only does this point to the need for caution in the interpretation of objective tests used as measures of achievement in educational psychology courses, but it also points to the need for more penetrating analyses of the psychological proficiencies of teachers-in-training both before and after instruction in educational psychology. Until such study is made it would seem difficult to set up standards of psychological proficiency for teachers that are realistic and valid.

2) The finding of a positive relationship between the attitude measure and the criterion tests suggests a host of problems concerning the rôle of values in psychology instruction. The fact that

close examination of both criterion tests revealed that the testee was required to make judgments not only of correctness of fact and general principles that account for human behaviors, but also of the correctness of certain procedures that depend upon value choices as well, suggests that educational psychologists need to attend more to the ways in which these value positions are learned. The question of when and how the 'should' questions are best handled in relation to the questions of 'what' and 'how' are in need of study.

3) The apparent lack of relationship between the amount of experience with children and the two criterion measures reinforces the conclusion that first-hand experience *per se* is not likely to enhance one's psychological understanding. Current emphasis in some teacher-education programs upon early field experiences may need to take into account more critically the types of guidance needed in such experiences if mature and valid understandings are to result.

BOOK REVIEWS

B. F. SKINNER. *Science and Human Behavior*. New York: The Macmillan Company, 1953, pp. 461. $4.00

Psychologists should have known what Skinner stands for from familiarity with his two previous books, *The Behavior of Organisms* and *Walden Two*, the former being a report of his investigations into the psychology of learning and the second being a novel which deals with the problem of Utopia. However, Skinner has really revealed his hand in this latest book, and has proclaimed himself to be a 'neo-behaviorist', taking on the mantle which John B. Watson let lapse over two decades ago. Skinner says in brief that behavior is the only datum concerning the mind with which science can deal and so he conceives of psychology strictly as the science of behavior. Skinner believes that the principal mission of science is to discover relationships that have a certain degree of generality and permanence. He feels that the only relationships with regard to behavior with which science can deal are those that connect behavior with the outer physical environment. Therefore, to Skinner behavior is the dependent variable and the environment is the independent variable, and the task of scientific psychology is to find the functional relationships between these two sets of variables.

Skinner has no use for so-called intervening variables—constructs which other psychologists have found necessary to set up in order to explain mental processes. A large part of this book is concerned with the endeavor to explain away the many intervening variables which have been proposed both in the common-sense world of everyday life and also by professional psychologists, and to reinterpret the phenomena which these intervening variables represent in terms of situation and response. Sometimes it is almost laughable to find Skinner taking an everyday concept which is accepted without difficulty by the common man and retranslate it into his own terms of reinforcement, extinction, discrimination and induction, What Skinner has done in the book is to set up a new psychology which will be consistent throughout in terms of a few principles which he believes to be basic to a psychology which is scientific.

But this book is more than a mere psychology, for in addition it has a program and a philosophy. Skinner says that "science is

more than a mere description of events. It is an attempt to discover order, to show that events stand in lawful relation to other events." (p.6) "Science not only describes, it predicts . . . nor is prediction the last word. To the extent that relevant conditions can be altered or otherwise controlled, the future can be controlled." This then is the keynote of the book. Skinner is interested in building a science of behavior in order to be able to control behavior.

After he finishes making his analysis of behavior, he turns to the problem of control. First, in Chapter 15, he describes self-control, and later, in Chapter 20, what he calls personal control and, in Chapter 21, group control. This is followed by a discussion of various controlling agencies in society, including government and law, religion, psychotherapy, economic control and education.

The final section of the book discusses the problem of human control in general, including such issues as, who shall control and for what purposes. Skinner belittles value judgments and believes that they have no place in the science of behavior. He then proceeds to propose his own basic value and make it the fundamental guiding principle on which all decisions with regard to control shall be made. In doing so he draws a parallel between the process of natural selection in the biological world and the process of selection of behavior in society. The guiding principle in biological selection is the survival of the species. In similar fashion he believes that the guiding principle in the selection of behavior is the survival of a culture. He says "a given culture is in short an experiment in behavior" (p 430) and he believes that the final criterion of behavior is the survival of a culture, and the problem of estimating the survival value of a culture is one for experimental determination. Skinner places greater value on the culture than on the individual, with all the social implications that such a point of view carries with it. He recognizes that this emphasis on the culture goes contrary to the emphasis in western democracy on the individual, but he believes that his thinking leads him inevitably to this conclusion.

Section II of the book is a summary of Skinner's views about learning, and anyone who wishes a compact statement of Skinner's point of view will find it in this section. Skinner, of course, was one of the first to propose a two-factor theory of learning, and this distinction is repeated in this book in terms of stimulus substitution

and response substitution. Skinner, however, does not take the extreme point of view which Mowrer has adopted in believing that these two kinds of learning are mutually exclusive. If this reviewer interprets Skinner correctly, he believes that most learning involves both stimulus substitution and response substitution. Skinner gives full recognition to the principle of secondary reinforcement which he calls 'conditioned reinforcement'.

He has a very interesting discussion of intermittent reinforcement and shows some of its practical implications. It will be remembered that some of the first experimental findings with regard to intermittent or partial reinforcement have come out of Skinner's researches. Skinner pays much attention to adversive behavior and avoidance behavior and there is a good discussion of punishment, drawing on contributions of Estes who did his researches under Skinner's supervision.

Skinner strives to be consistent throughout the book in avoiding the intervening variable, but when he comes to what he calls 'private events in a natural science' and the topic of the self, he occasionally slips from the rigorous position which he has attempted to adopt. Although he apparently is reporting Freud's point of view on page 284, he says "multiple selves or personalities are often said to be systematically related to each other. Freud conceived of the ego, superego, and id as distinguishable agents within the organism . . . the superego and the id were inevitably opposed to each other and Freud conceived of them as often in violent conflict. He appealed to a third agent—the ego—which, besides attempting to reach a compromise between the id and the superego, also dealt with the practical exigencies of the environment. We may quarrel with any analysis which appeals to a self or a personality as an inner determiner or action, but the facts which have been represented with such devices cannot be ignored. The three selves or personalities in the Freudian scheme represent important characteristics of behavior in a social milieu."

Apparently Skinner has made no effort to translate these intervening variables—the id, the superego and the ego—into the terms of his behavioristic categories but says "the facts which have been represented with such devices cannot be ignored."

This reviewer read with interest Skinner's discussion of the five controlling agencies in society, but feels that he has unfairly depicted at least two of them; namely, religion and education. Skin-

ner has described institutionalized religion and education as they now exist in current society, and has not given sufficient credit to religion and education which may be more of an inner, personal matter. In neither has he given sufficient credit to self-control, but has thought of both of these agencies as representing a fully externalized control by priests and teachers. As a matter of fact, because Skinner is so concerned with only relating behavior to external events he fails to give proper place to socalled inner or self-controls. In the discussion of self-control he rightly points out that all self-control is an outgrowth of external control. He fails, however, to give recognition to the fact that these external controls may have existed in the past and dropped out of the present picture so that for all intents and purposes these controls have been internalized or introjected. In omitting the recognition of these internalized controls as factors in religion and education he has given only a narrow and partial presentation of these controlling agencies.

Skinner seems to believe that control in education is very much a matter of punishment, and apparently does not recognize other types of control which are of a more positive nature. Evidence which this reviewer has gathered points conclusively to the fact that those teachers who depend most upon punishment as a means of securing control are among the least successful as teachers. This would point to the fact that there are other types of control in education to which Skinner has not given his attention.

The publishers of this book in their advertising refer to this as "an introductory psychology book." Certainly Skinner himself can be under no illusions that this is an introductory text. As a matter of fact, even the advanced student may find this book tough reading because it presupposes a rather thorough acquaintance with a vast amount of experimental material. This is a book without a statistical table and without the formal report of a single experiment. The unsophisticated reader may get the impression that Skinner has derived much of his material from the observation of and experimentation with pigeons because there are so many references to Skinner's later work with pigeons.

Any serious student of psychology, however, who wants to learn exactly what Skinner's system is and what applications he would

make of it to the practical world of affairs, will find this book a helpful reference for reading and study.

PERCIVAL M. SYMONDS

Teachers College, Columbia University

LESLIE PHILLIPS AND JOSEPH G. SMITH. *Rorschach Interpretation: Advanced Technique*. New York: Grune & Stratton, 1953, pp. 385.

Principles and procedures for Rorschach interpretation are nowhere more carefully explained than in this book. The authors set forth in detail many procedures clinical interpreters have other wise had to pick up by observation and imitation. The procedure is not thereby oversimplified or stereotyped. The authors firmly believe that every fact must be related to ever other, and they carry their suggestions to the utmost frontier of Rorschach interpretation. No cue is too faint, no relationship too speculative, to be included as a suggestion.

A unique feature of the book is that the authors are aware that their interpretations may be invalid. Their first paragraph is refreshing: "This book is intended primarily as a practical clinical reference. It is composed largely of statements about relationships between Rorschach performance and other behavior. Unfortunately, few of the relationships which are asserted have been corroborated; most represent guessed-at laws. They are perhaps best employed with equal parts of faith and skepticism." The writers warn that one proceeds by successive refinements, setting up tentative inferences on Card I which constitute predictions about subsequent cards; they can therefore be confirmed or modified (p. 261). Of course the hypothesis that results after ten cards then requires *its* "cross validation" from further data. It is in not stressing this need for testing the Rorschach interpretation that writers on the test err. Despite their sophistication and their excellent section on errors, Phillips and Smith do not succeed in maintaining their rigorous or skeptical attitude throughout the book.

Wherever the writers do reëxamine current doctrine, they make contributions. They often explain why interpretations are proposed. They develop much improved criteria for scoring location

and form level. They describe the interpretation of 'style' or non-scored behavior most helpfully. They present new developmental norms derived from theses by Hemmendinger, Siegel, and Friedman. The reviewer does not feel that their genetic approach to interpretation is always helpful. To say that an adult has a score common among children is not explained by describing this as a "genetically early level," or describing him as fixated at or regressed to this level. The significant point is that he has now acquired a positive way of behaving with his present organic potential, presumably because this way of behaving is relatively rewarding. Wherever analysis in genetic terms appears, it seems to crowd out discussion of learning processes or dynamics generally.

Recognizing the good features of the book, one still is startled by much of its content. "Shock on Card I is evidence for an unresolved and intense relation with a mother figure . . . Shock on Card II is characteristic of . . . uncontrolled destructive impulses." Remarks like "any time limits are characteristic of persons with obsessive-compulsive character features," "The concept 'teeth' . . . appears to be associated with masturbation and/or sibling rivalry as a reaction to frustrated dependency needs." Such comments are generally expanded in a way which gives them plausibility, yet the fact that the cables of evidence anchoring them to reality have all the tensile strength of moonbeams, and the authors' introductory warning that what they say may be either true or false, cause one to question just how such writings can have a place in a science.

If this book were read by a man from some 'ivory tower' who happened not to know that the Rorschach is currently taken seriously, it would sound to him like *Ralph 1245*C+ sounded to its first readers. Good science fiction is written with logic, any known facts, and imagination; at its best, its plausibilities foreshadow some things that science later establishes. Perhaps it is fairer to say that Rorschach interpretations are to the science of psychology as historical novels are to history. The writer seeks internal consistency and plausibility, with due regard for established fact. Where there is a gap in fact, the scholarly historian acknowledges the gap, but the fiction writer fills it with what might reasonably have happened, i.e., with an artistic creation.

The argument between scientist and clinician disappears if we see that clinical interpretation is an art at the present time. In the early days of testing it was a technology, wrong at times but never

seeking to be creative or intuitive. Now, instead, the clinician is first of all a portraitist who abhors unfilled areas on the canvas. This occurs because there is a practical demand for answers about people, not for uncertainties. Society has rewarded this new species of artist; we can scarcely criticize him for satisfying the demand. New entrants in the field want to be successful artists, and books like the present one which tell how to complete an attractive and convincing portrait are wanted. When authors like Phillips and Smith can also help keep portraits closer to reality than they would otherwise be, that is so much gained. Once we grant that the clinician is going to provide descriptions which go beyond proven fact, we should welcome any book which contributes to making him a good non-scientist.

The scientist cannot complain if the portrait is used as a guide to further inquiry about the patient or about personality in general. Both science fiction and historical fiction have suggested leads to serious investigations of their fields. This emphatically demands that the clinician's suggested interpretations should send further study in promising directions, not up 'blind alleys'. There is good reason to think that the Rorschach or other clinical interviewing methods can do this often enough to be useful, even though, according to the evidence, their suggestions on any case are not correct enough to be trusted without further support.

The scientist should object when he is told that he or the practitioner who receives a Rorschach analysis should believe it. Among all the books yet written on the Rorschach, the book under review comes closest to filling a practical demand for training clinicians to be artistic portraitists, without making unacceptable claims that such portraits are dependable.

LEE J. CRONBACH
University of Illinois

ERNEST GARDNER. *Fundamentals of Neurology*. Philadelphia: W. B. Saunders Company, 1952, pp. 359.

Neurology is, in this second edition of the textbook on the subject as in the first edition, presented as a series of basic concepts—revised in the light of growing knowledge in the field. The material selected for inclusion is based upon dissection and microscopic studies of the nervous system, experimental data, and the study of

neurological disorders. In the nineteen chapters of this small volume are included a consideration of functional as well as structural topics. Also included is a glossary of the many new terms of Greek and Latin sources, and an index. Each chapter ends with a brief biographical sketch of the people who have contributed to the topics considered in the chapter.

Psychologists will be interested in this text on neurology as such and also in the treatment of the problems of learning and language mechanisms in the last chapter called "Associative Functions of the Cerebral Cortex". Treatment of learning and conditioned responses are presented in a very compact manner, and the frame of reference for the consideration can be gathered from the psychological sources used by the author. These include Hilgard and Marquis' *Conditioning and Learning*, Lashley's *Brain Mechanisms and Intelligence*, McGeogh's *Psychology of Human Learning*, Murchison's *Handbook of General Experimental Psychology*, Pavlov's *Conditioned Reflexes*, and Young's *Emotion in Man and Animal.*

Psychological Service Center H. MELTZER